

HAUPTAUFSÄTZE

Verzerrungstensor, Verzerrungsdeviator und Spannungstensor bei endlichen Formänderungen

Von Hans Richter in Haltingen/Lörrach

Es werden Postulate aufgestellt, denen bei der Bildung des Verzerrungstensors, des Verzerrungsdeviators und des Spannungstensors zu genügen ist, und hieraus die allgemeine Gestalt dieser Tensoren in beliebigen Koordinaten abgeleitet. Als einfachste Definition des Verzerrungstensors erscheint die gemischt-variante logarithmische Deformationsmatrix, wo der Deviator in üblicher Weise gebildet werden kann, und wo die Invarianten des letzteren die Beanspruchung invariant charakterisieren. Bei entsprechender Definition des Spannungstensors bleibt die Gestalt des allgemeinen Elastizitätsgesetzes invariant gegen Koordinatentransformation.

Postulates are laid down that have to be satisfied on forming the distortion tensor, the distortion deviator, and the tension tensor, and thus the general form of these tensors are deduced in arbitrary co-ordinates. The combined-variant logarithmical deformation matrix proves the simplest definition of the distortion tensor. The deviator may be formed in the usual manner, and the invariants of it characterize the strain in an invariant way. If the tension tensor is defined accordingly, the form of the general law of elasticity continues to be invariant to co-ordinate transformations.

On établit des postulats pour la formation du tenseur de déformation, du déviateur de déformation et du tenseur de tension. La forme générale de ces tenseurs en coordonnées arbitraires en est déduite. La matrice logarithmique (mixte-variante) de déformation fournit la plus simple définition du tenseur de déformation. Le déviateur peut être formé comme de coutume et ses invariants caractérisent la sollicitation d'une manière invariante. Le tenseur de tension étant défini conformément, la forme de la loi générale d'élasticité reste invariante dans toute transformation de coordonnées.

Устанавливаются постулаты, из которых нужно исходить при построении тензора деформаций, девиатора деформаций и тензора напряжений. Отсюда определяется общая форма этих тензоров в произвольных координатах. Наиболее простым определением тензора деформаций представляется смешанно-вариантная логарифмическая матрица деформаций, причем девиатор строится обычным способом. Инварианты девиатора инвариантно характеризуют напряжения. При соответственном определении тензора напряжений форма основных уравнений теории упругости инвариантна по отношению к преобразованию координат.

§ 1. Einleitung

In der Theorie endlicher elastischer oder plastischer Verformungen geht man im allgemeinen von demjenigen Verzerrungstensor aus, der für allgemeine Koordinaten durch Bildung der Differenz der Quadrate der Linienelemente im deformierten und im Ausgangszustand entsteht¹⁾. Die Verwendung gerade dieser Charakterisierung des Verzerrungszustandes ist natürlich nicht zwingend vorgeschrieben. Im Gegenteil erscheint bei einer genaueren Analyse, die im folgenden durchgeführt werden soll, gerade diese übliche Definition des Verzerrungstensors nicht als diejenige, die dem Problem der Untersuchung endlicher Verzerrungen besonders gut angepaßt ist. So führt bereits die Aufgabe, aus dem üblichen Verzerrungstensor einen Deviator abzuleiten, der in Absonderung der Volumänderung nur die Gestaltänderung charakterisiert, zu eigenartigen Schwierigkeiten und Mehrdeutigkeiten¹⁾. Der tiefere Grund hierfür liegt wohl darin, daß man sich bei der Behandlung der endlichen Formänderungen zu stark an das Vorbild der infinitesimalen Verzerrungen gehalten hat, bei denen man eine beliebige Deformation durch additive Aufspaltung in den symmetrischen und den schiefssymmetrischen Anteil in eine reine Streckung und eine reine Drehung zerlegen kann. Bei endlichen Verformungen ist diese additive Aufspaltung jedoch nicht mehr möglich; an ihre Stelle tritt eine multiplikative Zerlegung der allgemeinen Deformation in eine Drehung und eine Streckung, wobei diese Faktoren nicht mehr kommutativ sind. Jeder Versuch, Definitionen durch additive Zerspaltung zu bilden, muß daher auf grundsätzliche Schwierigkeiten stoßen.

Wir wollen nun in dieser Arbeit so vorgehen, daß wir — gewissermaßen axiomatisch — an die zu bildenden Definitionen von vornherein gewisse Forderungen stellen, die uns als zweckmäßig erscheinen, und dann zeigen, wie aus diesen Forderungen sich bestimmte Möglichkeiten für diese Definitionen als besonders naheliegend ergeben.

¹⁾ Vgl. R. Moufang: Volumtreue Verzerrungen bei endlichen Formänderungen. Z. angew. Math. Mech., Bd. 25/27 (1947), S. 209—214.

§ 2. Bezeichnungen und Hilfssätze

a) Bezeichnungen

1. Mit großen gotischen Buchstaben \mathfrak{A} , \mathfrak{B} , ... bezeichnen wir dreireihige quadratische Matrizen²⁾, $a_{ik} = (\mathfrak{A})_{ik}$ ist das Element in der i -ten Zeile und der k -ten Spalte. $|\mathfrak{A}|$ ist die Determinante von \mathfrak{A} . $\{\mathfrak{A}\}$ ist die Spur von \mathfrak{A} , also die Summe der Elemente der Hauptdiagonale. $\bar{\mathfrak{A}}$ ist die an der Hauptdiagonale Gespiegelte von \mathfrak{A} . \mathbb{E} ist die Einheitsmatrix. \mathfrak{A}^{-1} ist die Inverse zu \mathfrak{A} .

2. Kleine gotische Buchstaben \mathfrak{x} , \mathfrak{y} , ... bedeuten Vektoren: $\mathfrak{x} = (x_1, x_2, x_3)$; $x_p = (\mathfrak{x})_p$. $\mathfrak{x} \cdot \mathfrak{y}$ ist das innere Produkt. $\mathfrak{x} \times \mathfrak{y}$ ist das äußere Produkt.

3. $\mathfrak{A}\mathfrak{x}$ entsteht durch Anwendung von \mathfrak{A} auf \mathfrak{x} : $(\mathfrak{A}\mathfrak{x})_i = \sum_k a_{ik} x_k$.

4. Produkte $\mathfrak{B}\mathfrak{A}$ sind von rechts nach links zu lesen: $(\mathfrak{B}\mathfrak{A})\mathfrak{x} = \mathfrak{B}(\mathfrak{A}\mathfrak{x})$.

5. Ist $f(x) = \sum b_n \cdot x^n$, so ist unter Voraussetzung der Konvergenz: $f(\mathfrak{A}) = \sum b_n \mathfrak{A}^n$; $df(\mathfrak{A}) = f(\mathfrak{A} + d\mathfrak{A}) - f(\mathfrak{A})$, was mit $f'(\mathfrak{A})d\mathfrak{A}$ nur bei $\mathfrak{A} \cdot d\mathfrak{A} = d\mathfrak{A} \cdot \mathfrak{A}$ übereinstimmt³⁾.

b) Hilfssätze

(2.1) $\{\mathfrak{A}_1 \mathfrak{A}_2 \dots \mathfrak{A}_n\}$ bleibt bei zyklischer Vertauschung der Faktoren ungeändert.

(2.2) Alle Invarianten von \mathfrak{A} gegen affine Transformation $\mathfrak{A} \rightarrow \mathbb{C}\mathfrak{A}\mathbb{C}^{-1}$ sind Funktionen der drei Invarianten $j = \{\mathfrak{A}\}$, $k = \{\mathfrak{A}^2\}$ und $l = \{\mathfrak{A}^3\}$. Die charakteristische Gleichung von \mathfrak{A} ist:

$$x^3 - j \cdot x^2 + \frac{1}{2}(j^2 - k)x - \left(\frac{1}{3}l - \frac{1}{2}jk + \frac{1}{6}j^3\right) = 0.$$

(2.3) Es ist $f(\mathbb{C}\mathfrak{A}\mathbb{C}^{-1}) = \mathbb{C}f(\mathfrak{A})\mathbb{C}^{-1}$.

(2.4) Hat \mathfrak{A} positiv reelle Eigenwerte, so ist $\ln \mathfrak{A}$ definiert. Es ist $\{\ln \mathfrak{A}\} = \ln |\mathfrak{A}|$.

(2.5) Es ist $\{\mathfrak{B} df(\mathfrak{A})\} = \{\mathfrak{B} f'(\mathfrak{A}) d\mathfrak{A}\}$, falls $\mathfrak{B}\mathfrak{A} = \mathfrak{A}\mathfrak{B}$, jedoch nicht notwendig $\mathfrak{B} \cdot d\mathfrak{A} = d\mathfrak{A} \cdot \mathfrak{B}$ ist.

(2.6) In kartesischen Koordinaten ist eine reine Streckung \mathbb{S} symmetrisch mit positiven Eigenwerten.

(2.7) In kartesischen Koordinaten gilt für eine euklidische Transformation \mathfrak{R} : $\mathfrak{R}\bar{\mathfrak{R}} = \mathbb{E}$.

(2.8) Ein beliebiges \mathfrak{A} mit $|\mathfrak{A}| \neq 0$ läßt sich eindeutig darstellen in der Form $\mathfrak{A} = \mathbb{S} \cdot \mathfrak{R}$, d. h. als euklidische Transformation mit nachfolgender reiner Streckung. Bei $|\mathfrak{A}| > 0$ ist \mathfrak{R} eine direkte Transformation, also eine reine euklidische Drehung.

(2.9) Es ist $\mathfrak{x} \cdot \mathfrak{A}\mathfrak{y} = \mathfrak{y} \cdot \bar{\mathfrak{A}}\mathfrak{x}$.

(2.10) Es sei $\mathfrak{y} = \mathfrak{U}\mathfrak{x}$ eine Koordinatentransformation, bei der \mathfrak{A} in \mathfrak{A}^* übergeht. \mathfrak{A} ist ein

zweifach kontravarianter Tensor, falls $\mathfrak{A}^* = \mathfrak{U}\mathfrak{A}\bar{\mathfrak{U}} \cdot |\mathfrak{U}|^n$,

zweifach kovarianter Tensor, falls $\mathfrak{A}^* = \bar{\mathfrak{U}}^{-1} \mathfrak{A} \mathfrak{U}^{-1} \cdot |\mathfrak{U}|^n$,

kontravariant-kovarianter Tensor, falls $\mathfrak{A}^* = \mathfrak{U}\mathfrak{A}\mathfrak{U}^{-1} \cdot |\mathfrak{U}|^n$,

kovariant-kontravarianter Tensor, falls $\mathfrak{A}^* = \bar{\mathfrak{U}}^{-1} \mathfrak{A} \bar{\mathfrak{U}} \cdot |\mathfrak{U}|^n$ ist.

Bei $n = 0$ ist \mathfrak{A} ein eigentlicher Tensor; bei $n \neq 0$ eine tensorielle Dichte. (Die Übereinstimmung dieser etwas weniger geläufigen Darstellung der Tensoreigenschaft mit der üblichen ergibt sich unmittelbar, wenn man symbolisch $(\mathfrak{A})_{ik} = x_i y_k$ setzt, wo \mathfrak{x} und \mathfrak{y} kontravariante oder kovariante Vektoren sind).

(2.11) Es sei $\mathfrak{x}' = \mathfrak{U}\mathfrak{x}$ und $\mathfrak{y}' = \mathfrak{U}\mathfrak{y}$; dann ist $\mathfrak{x}' \times \mathfrak{y}' = |\mathfrak{U}| \cdot \bar{\mathfrak{U}}^{-1} (\mathfrak{x} \times \mathfrak{y})$.

§ 3. Der Verzerrungstensor

Wir wollen uns nun überlegen, welche Forderungen wir billigerweise an die Verzerrungsmatrix stellen können, um dann die Realisierbarkeit dieser Forderungen zu untersuchen.

Es sei \mathfrak{A} die Matrix, die die Umgebung eines Punktes $\hat{\mathfrak{x}}$ auf die Umgebung des Bildpunktes \mathfrak{x} abbildet:

$$(3.1) \quad d\mathfrak{x} = \mathfrak{A} d\hat{\mathfrak{x}}.$$

²⁾ Ob eine Matrix ein Tensor ist, ist an (2.10) zu sehen.

³⁾ Vgl. jedoch (2.5).

\mathfrak{A} ist die Funktionalmatrix

$$(3.2) \quad (\mathfrak{A})_{ik} = \frac{\partial x_i}{\partial \hat{x}_k}, \quad |\mathfrak{A}| > 0$$

und gibt den erreichten Verzerrungszustand an. Bei plastischem Material, wo der Spannungszustand nicht nur vom erreichten Verzerrungszustand abhängt, sondern auch von dem Wege, der zu diesem führte, ist dann die Angabe von \mathfrak{A} allein nicht ausreichend. Bei elastischem Material dagegen genügt \mathfrak{A} zur Charakterisierung der Verzerrung. Bei anisotropem Material ist auch die in \mathfrak{A} enthaltene Drehung wesentlich. Es ist dann \mathfrak{A} selbst zur Beschreibung der Verzerrung zu verwenden, während alle Verzerrungstensoren, die wie der gewöhnliche eine euklidische Drehung eliminieren, unbrauchbar sind. Die Aufstellung von solchen Verzerrungstensoren hat daher überhaupt nur für isotropes Material Sinn.

a) Postulate

D mgemäß soll nun unter der ausdrücklichen Voraussetzung der Anwendbarkeit auf isotropes Material zu \mathfrak{A} eine Verzerrungsmatrix $\mathfrak{B}(\mathfrak{A})$ definiert werden⁴⁾. Während nun \mathfrak{A} sicher kein Tensor ist, da \mathfrak{A} auf zwei verschiedene Punkte bezogen ist, wollen wir die Tensor-eigenschaft von \mathfrak{B} fordern. Damit ergibt sich das erste Postulat:

V1: \mathfrak{B} ist ein Tensor, der aus \mathfrak{A} und den Matrizen der Metrik in \hat{x} und x gebildet werden kann.

Weit rhin soll bei \mathfrak{B} die unwesentliche in \mathfrak{A} enthaltene Drehung unberücksichtigt bleiben. Es soll sich also \mathfrak{B} nicht ändern, wenn v o r Ausübung von \mathfrak{A} erst eine euklidische Drehung \mathfrak{R} in \hat{x} durchgeführt wird. Man kann sich statt dessen auch auf den Standpunkt stellen, daß eine n a c h \mathfrak{A} in x durchgeführte Drehung den Verzerrungstensor nicht beeinflussen soll. Dies würde bedeuten, daß \mathfrak{A} als Verzerrung in \hat{x} mit nachfolgender unwesentlicher Drehung aufgefaßt wird. Den so zu \hat{x} gehörigen Verzerrungstensor wollen wir mit $\hat{\mathfrak{B}}$ bezeichnen. Die Analyse von \mathfrak{B} und $\hat{\mathfrak{B}}$ ist vollkommen analog, so daß wir uns im folgenden auf die von \mathfrak{B} beschränken und die für $\hat{\mathfrak{B}}$ analogen Ergebnisse lediglich vermerken, wobei die entsprechenden Größen durch ein \wedge gekennzeichnet werden.

Die genannte Eigenschaft von \mathfrak{B} und $\hat{\mathfrak{B}}$ drückt sich nun aus durch das Postulat

V2: $\mathfrak{B}(\mathfrak{A} \mathfrak{R}) = \mathfrak{B}(\mathfrak{A})$, resp. $\hat{\mathfrak{B}}(\mathfrak{R} \mathfrak{A}) = \hat{\mathfrak{B}}(\mathfrak{A})$.

Weiterhin verlangen wir noch ein Superpositionsprinzip für koaxiale reine Streckungen durch Postulat

V3: Es seien \mathfrak{S}_1 und \mathfrak{S}_2 zwei koaxiale Streckungen: $\mathfrak{S}_1 \mathfrak{S}_2 = \mathfrak{S}_2 \mathfrak{S}_1$. Es sei $\mathfrak{B}_1 = \mathfrak{B}(\mathfrak{S}_1)$, $\mathfrak{B}_2 = \mathfrak{B}(\mathfrak{S}_2)$ und $\mathfrak{B} = \mathfrak{B}(\mathfrak{S}_1 \mathfrak{S}_2)$. Dann soll es eine umkehrbar eindeutige Funktion $f(x)$ geben, so daß $f(\mathfrak{B}_1) + f(\mathfrak{B}_2) = f(\mathfrak{B})$ ist. f kann gegebenenfalls abhängig vom Koordinatensystem sein.

Schließlich müssen wir noch verlangen, daß für infinitesimale Verzerrungen die neue Definition in die alte übergeht. Dies liefert die Limesbeziehung

V4: Für infinitesimale Deformationen $\mathfrak{E} + d\mathfrak{B}$ in kartesischen Koordinaten geht der Verzerrungstensor in $\frac{1}{2}(d\mathfrak{B} + \overline{d\mathfrak{B}}) + o(d\mathfrak{B})$ über⁵⁾.

b) Die Realisierung der Postulate in kartesischen Koordinaten

Aus Vereinfachungsgründen wollen wir zunächst kartesische Koordinaten voraussetzen. Original- und Bildpunkt der Deformation nennen wir dann $\hat{\eta}$ und η . Die Deformationsmatrix heiße jetzt \mathfrak{B} . Die zugehörigen Verzerrungstensoren sind \mathfrak{B} und $\hat{\mathfrak{B}}$.

Gemäß (2.8) schreiben wir zunächst

$$(3.3) \quad \mathfrak{B} = \mathfrak{C} \mathfrak{R} = \mathfrak{R} \hat{\mathfrak{C}} \text{ bei } \hat{\mathfrak{C}} = \mathfrak{R}^{-1} \mathfrak{C} \mathfrak{R}.$$

Zur Auffindung dieser Zerlegung bildet man zunächst $\mathfrak{B} \overline{\mathfrak{B}}$. Nun ist für $x \neq 0$: $0 < \overline{\mathfrak{B}} x \cdot \overline{\mathfrak{B}} x$, was mit Hilfe von (2.9) liefert: $0 < x \cdot \mathfrak{B} \mathfrak{B} x$. Die symmetrische Matrix $\mathfrak{B} \overline{\mathfrak{B}}$ ist also positiv

⁴⁾ Die Bezeichnung $\mathfrak{B}(\mathfrak{A})$ soll dabei nicht heißen, daß \mathfrak{B} eine Funktion von \mathfrak{A} im Sinne von § 2. Pkt. 5 ist, sondern nur, daß \mathfrak{B} zu \mathfrak{A} gehört.

⁵⁾ Wie üblich bedeutet $y = o(x)$: $\lim_{x \rightarrow 0} \frac{y}{x} = 0$.

definit und hat daher eindeutig eine positiv definite Quadratwurzel $\mathfrak{S} = \sqrt{\mathfrak{B} \mathfrak{B}}$. \mathfrak{R} ergibt sich dann als $\mathfrak{R} = \mathfrak{S}^{-1} \mathfrak{B}$. Entsprechend ist $\hat{\mathfrak{E}}^2 = \bar{\mathfrak{B}} \mathfrak{B}$.

Nach V 2 ist $\mathfrak{B}(\mathfrak{B}) = \mathfrak{B}(\mathfrak{S})$, resp. $\hat{\mathfrak{B}}(\mathfrak{B}) = \hat{\mathfrak{B}}(\hat{\mathfrak{E}})$. Wir können uns also auf Verzerrungstensoren beschränken, die zu reinen Streckungen gehören.

Es sei nun \mathfrak{S} eine infinitesimale Streckung: $\mathfrak{S} = \mathfrak{E} + d\mathfrak{S}$. Dann ist nach V 4: $\mathfrak{B}(\mathfrak{E} + d\mathfrak{S}) = d\mathfrak{S} + o(d\mathfrak{S})$ und $\mathfrak{B}(\mathfrak{E} + \lambda d\mathfrak{S}) = \lambda d\mathfrak{S} + o(d\mathfrak{S})$, wenn λ eine positive Zahl ist. Aus V 3 ergibt sich dann $f(d\mathfrak{S} + o(d\mathfrak{S})) + f(\lambda d\mathfrak{S} + o(d\mathfrak{S})) = f((1 + \lambda) d\mathfrak{S} + o(d\mathfrak{S}))$. Da dies für jedes λ und $d\mathfrak{S}$ gelten muß, folgt für kleine x : $f(x) = x + o(x)^6$. Setzen wir nun $\mathfrak{B} = f(\mathfrak{B})$, so wird damit für infinitesimale Streckungen: $\mathfrak{B}(\mathfrak{E} + d\mathfrak{S}) = d\mathfrak{S} + o(d\mathfrak{S})$.

Sei nun wieder \mathfrak{S} eine endliche reine Streckung, so läßt sich wegen der positiven Eigenwerte von \mathfrak{S} bilden:

$$(3.4) \quad \mathfrak{Q} = \ln \mathfrak{S}; \quad \text{resp.} \quad \hat{\mathfrak{Q}} = \ln \hat{\mathfrak{E}}: \text{„logarithmische Deformationsmatrix“.}$$

Es ist dann: $\frac{1}{n} \mathfrak{Q} = \ln \sqrt[n]{\mathfrak{S}}$ und damit für große n : $\sqrt[n]{\mathfrak{S}} = \mathfrak{E} + \frac{1}{n} \mathfrak{Q} + o\left(\frac{1}{n}\right)$. Also wird $\mathfrak{B}(\sqrt[n]{\mathfrak{S}}) = \frac{1}{n} \mathfrak{Q} + o\left(\frac{1}{n}\right)$. Nach V 3 ist weiter $\mathfrak{B}(\mathfrak{S}) = n \cdot \mathfrak{B}(\sqrt[n]{\mathfrak{S}}) = \mathfrak{Q} + n \cdot o\left(\frac{1}{n}\right)$. Da die linke Seite dieser Gleichung von n unabhängig ist, können wir n gegen unendlich gehen lassen und erhalten: $\mathfrak{B}(\mathfrak{S}) = \mathfrak{Q}$. Hieraus folgt insbesondere, daß $f(x)$ bis auf einen willkürlichen Faktor eindeutig bestimmt ist.

Ist nun g die Umkehrfunktion zu f , so haben wir schließlich, da ja \mathfrak{Q} eine umkehrbar eindeutige Funktion von \mathfrak{S} und damit auch von $\mathfrak{S}^2 = \mathfrak{B} \mathfrak{B}$ ist:

$$(3.5) \quad \mathfrak{B} = g(\mathfrak{Q}) = h(\mathfrak{S}) = k(\mathfrak{B} \mathfrak{B}); \quad \text{resp.} \quad \hat{\mathfrak{B}} = g(\hat{\mathfrak{Q}}) = h(\hat{\mathfrak{E}}) = k(\bar{\mathfrak{B}} \mathfrak{B}).$$

Der Ansatz (3.5) genügt umgekehrt stets den Postulaten V 2 und V 3, wobei f eindeutig als Umkehrfunktion von g zu wählen ist, während die Erfüllung der Limesbedingung V 4 fordert, daß für kleine x gilt:

$$(3.5^a) \quad g(x) = x + o(x); \quad h(1+x) = x + o(x); \quad k(1+x) = \frac{1}{2} x + o(x).$$

In der Tat ist dann für infinitesimale Verzerrungen $\mathfrak{B} = \mathfrak{E} + d\mathfrak{B}$:

$$\begin{aligned} \mathfrak{B} \bar{\mathfrak{B}} &= \mathfrak{E} + (d\mathfrak{B} + d\bar{\mathfrak{B}}) + o(d\mathfrak{B}), \quad \mathfrak{S} = \sqrt{\mathfrak{B} \bar{\mathfrak{B}}} = \mathfrak{E} + \frac{1}{2} (d\mathfrak{B} + d\bar{\mathfrak{B}}) + o(d\mathfrak{B}), \\ \mathfrak{Q} = \ln \mathfrak{S} &= \frac{1}{2} (d\mathfrak{B} + d\bar{\mathfrak{B}}) + o(d\mathfrak{B}). \end{aligned}$$

Es ist also bei jedem (3.5^a) genügenden g : $\mathfrak{B} = \frac{1}{2} (d\mathfrak{B} + d\bar{\mathfrak{B}}) + o(d\mathfrak{B})$.

Jeder mit unseren Postulaten verträgliche Verzerrungstensor ist damit als Funktion der logarithmischen Deformationsmatrix erkannt. Vom Standpunkte unserer Postulate aus erscheint $\mathfrak{B} = \mathfrak{Q}$ als die einfachste Definition der Verzerrungsmatrix, da hier das Superpositionsprinzip mit $f(x) \equiv x$ erfüllt ist. Wir werden später sehen, daß auch für die Bildung des Deviators diese Definition als die einfachste erscheinen wird.

Wird anschließend an \mathfrak{R} noch eine euklidische Drehung \mathfrak{R}_1 durchgeführt, so geht wegen $\mathfrak{R}_1 \mathfrak{R} = \mathfrak{R}_1 \mathfrak{S} \mathfrak{R} = \mathfrak{R}_1 \mathfrak{S} \mathfrak{R}_1^{-1} \cdot \mathfrak{R}_1 \mathfrak{R}$ die Streckung \mathfrak{S} in $\mathfrak{R}_1 \mathfrak{S} \mathfrak{R}_1^{-1}$ über. Den gleichen Übergang haben wir, wenn wir eine euklidische Koordinatentransformation $\eta_1 = \mathfrak{R}_1 \eta$ durchführen. Gemäß (2.3) geht dann \mathfrak{B} in $h(\mathfrak{R}_1 \mathfrak{S} \mathfrak{R}_1^{-1}) = \mathfrak{R}_1 \mathfrak{B} \mathfrak{R}_1^{-1}$ über. Die Achsen von \mathfrak{B} werden also bei nachträglicher Anwendung von \mathfrak{R}_1 einfach mitgedreht. Bei der anderen Auffassung der letzten Formel als Ergebnis einer Koordinatentransformation sehen wir an (2.10), daß \mathfrak{B} sich wie ein Tensor transformiert, wobei wegen $\mathfrak{R}_1^{-1} = \bar{\mathfrak{R}}_1$ kein Unterschied bezüglich der Varianz vorhanden ist. Entsprechendes gilt natürlich für $\bar{\mathfrak{B}}$.

c) Erweiterung auf krummlinige Koordinaten

Wir gehen jetzt von den kartesischen Koordinaten η auf beliebige Koordinaten \mathfrak{x} über: $\mathfrak{x} = \mathfrak{x}(\eta)$. Für eine Umgebung des undeformierten Materials sei $d\mathfrak{x} = \hat{\mathfrak{U}} d\eta$, für die entsprechende Umgebung im deformierten Material sei $d\mathfrak{x} = \mathfrak{U} d\eta$. $\hat{\mathfrak{U}}$ und \mathfrak{U} sind die Funktionalmatrizen von $\mathfrak{x} = \mathfrak{x}(\eta)$ in $\hat{\mathfrak{x}}$ und \mathfrak{x} .

⁶⁾ Da mit f jedes Vielfache von f dem Postulat V 3 genügt, kann man $f'(0)$ zu 1 normieren.

Für ein Linienelement in \mathfrak{x} erhalten wir mit Hilfe von (2.9): $d\mathfrak{x} = \mathfrak{U}^{-1} d\mathfrak{x} \cdot \mathfrak{U}^{-1} d\mathfrak{x} = d\mathfrak{x} \cdot \mathfrak{U}^{-1} \mathfrak{U}^{-1} d\mathfrak{x}$. Also ist

$$(3.6^a) \quad \mathfrak{G} = \overline{\mathfrak{G}} = (\mathfrak{U} \overline{\mathfrak{U}})^{-1}$$

die Matrix der Metrik in \mathfrak{x} . Entsprechend definiert

$$(3.6^b) \quad \hat{\mathfrak{G}} = (\hat{\mathfrak{U}} \hat{\mathfrak{U}})^{-1}$$

die Metrik in $\hat{\mathfrak{x}}$.

Die Deformation des Materials erscheint jetzt als: $d\mathfrak{x} = \mathfrak{U} \mathfrak{B} \hat{\mathfrak{U}}^{-1} d\hat{\mathfrak{x}}$. \mathfrak{B} ist also übergegangen in

$$(3.7) \quad \mathfrak{A} = \mathfrak{U} \mathfrak{B} \hat{\mathfrak{U}}^{-1}, \quad (\mathfrak{A})_{ik} = \frac{\partial x_i}{\partial \hat{x}_k}$$

Umgekehrt ist

$$(3.7^*) \quad \mathfrak{B} = \mathfrak{U}^{-1} \mathfrak{A} \hat{\mathfrak{U}} \quad \text{und} \quad \overline{\mathfrak{B}} = \overline{\mathfrak{U}}^{-1} \overline{\mathfrak{A}} \overline{\hat{\mathfrak{U}}}^{-1},$$

woraus wir sofort erhalten:

$$(3.8) \quad \begin{cases} \mathfrak{S}^2 = \mathfrak{B} \overline{\mathfrak{B}} = \mathfrak{U}^{-1} \mathfrak{A} \overline{\mathfrak{A}}^{-1} \overline{\mathfrak{U}}^{-1} \\ \hat{\mathfrak{S}}^2 = \overline{\mathfrak{B}} \mathfrak{B} = \overline{\mathfrak{U}}^{-1} \overline{\mathfrak{A}} \mathfrak{A} \hat{\mathfrak{U}} \end{cases}$$

Mit Hilfe der beiden letzten Formeln können wir also die zu den kartesischen Koordinaten gehörigen Matrizen \mathfrak{B} , \mathfrak{S} und $\hat{\mathfrak{S}}$ durch \mathfrak{A} und die Übergangsmatrizen \mathfrak{U} und $\hat{\mathfrak{U}}$ ausdrücken.

α) Fall des ungemischten Tensors

Wir nehmen zunächst an, es sei der Verzerrungstensor \mathfrak{B} zweifach kontravariant definiert und genüge den Postulaten $V1$ bis $V4$. Dann muß nach (2.10) sein:

$$\mathfrak{B} = \mathfrak{U} \mathfrak{B} \overline{\mathfrak{U}},$$

wobei \mathfrak{B} einer der Tensoren (3.5) ist.

Um die besondere Gestalt von \mathfrak{B} kennenzulernen, nehmen wir den Spezialfall, daß \mathfrak{B} eine reine Streckung \mathfrak{S} in den Koordinatenachsen und \mathfrak{U} damit koaxial ist, also

$$\mathfrak{S} = \begin{pmatrix} \lambda_1 & 0 \\ & \lambda_2 \\ 0 & & \lambda_3 \end{pmatrix} \quad \text{und} \quad \mathfrak{U} = \begin{pmatrix} \varrho_1 & 0 \\ & \varrho_2 \\ 0 & & \varrho_3 \end{pmatrix}.$$

Dann ist wegen (3.5) \mathfrak{B} ebenfalls auf Hauptachsen mit den Eigenwerten $\varrho_v^2 \cdot h(\lambda_v)$. Das Superpositionsprinzip fordert nun die Existenz einer Funktion $f(x)$, deren Koeffizienten die ϱ_v enthalten können, so daß gilt:

$$(3.9) \quad f(\varrho_v^2 h(\lambda_v)) + f(\varrho_v^2 h(\mu_v)) = f(\varrho_v^2 h(\lambda_v \mu_v))$$

für beliebige λ_v und μ_v . Es ist daher

$$f(\varrho_v^2 h(\lambda)) = C_v \cdot \ln \lambda, \quad C_v = C_v(\varrho_1, \varrho_2, \varrho_3).$$

Durch Differentiation nach λ erhalten wir hieraus:

$$(3.10) \quad \varrho_v^2 f'(\varrho_v^2 h(\lambda)) \cdot h'(\lambda) = C_v \cdot \frac{1}{\lambda}.$$

Setzen wir speziell $\lambda = 1$, so wird wegen (3.5^a): $\varrho_v^2 \cdot f'(0) = C_v$, so daß bei der Normierung $f'(0) = 1$ wird:

$$f'(\varrho_v^2 h(\lambda)) = \frac{1}{\lambda \cdot h'(\lambda)}.$$

Die rechte Seite dieser Gleichung ist unabhängig von den ϱ_v . Es muß daher $f'(x) \equiv f'(0) = 1$ und damit $h(\lambda) = \ln \lambda$ sein. Es ist dann $\mathfrak{B} = \ln \mathfrak{S} = \mathfrak{S}$ und damit

$$\mathfrak{B} = \mathfrak{U} \mathfrak{S} \overline{\mathfrak{U}}$$

zu setzen. Umgekehrt erfüllt diese Definition für \mathfrak{B} offenbar auch bei beliebiger Wahl von \mathfrak{B} und \mathfrak{U} alle Postulate $V1$ bis $V4$, da ja für \mathfrak{S} das Superpositionsprinzip rein additiv ist und sich daher bei Multiplikation von links mit \mathfrak{U} und von rechts mit $\overline{\mathfrak{U}}$ einfach auf ein additives Gesetz für \mathfrak{B} überträgt.

Es gibt also für ungemischt tensorielles \mathfrak{B} überhaupt nur eine mit unseren Postulaten verträgliche Definitionsmöglichkeit: $\mathfrak{B} = \mathfrak{U} \mathfrak{Q} \mathfrak{U}$.

Wegen $\mathfrak{Q} = \ln \mathfrak{S} = \frac{1}{2} \ln \mathfrak{S}^2$ erhalten wir schließlich nach (3.8) den Ausdruck:

$$(3.11) \quad \mathfrak{B} = \frac{1}{2} \mathfrak{U} \ln (\mathfrak{U}^{-1} \mathfrak{A} \hat{\mathfrak{G}}^{-1} \overline{\mathfrak{A}} \overline{\mathfrak{U}}^{-1}) \mathfrak{U}.$$

Entsprechend ergibt sich

$$\hat{\mathfrak{B}} = \frac{1}{2} \hat{\mathfrak{U}} \ln (\hat{\mathfrak{U}} \overline{\mathfrak{A}} \mathfrak{G} \mathfrak{A} \hat{\mathfrak{U}}) \hat{\mathfrak{U}}.$$

Entwickeln wir den \ln für nicht zu große Verzerrungen in eine Potenzreihe, so sehen wir, daß in der Tat \mathfrak{B} und $\hat{\mathfrak{B}}$ nur von \mathfrak{A} , $\hat{\mathfrak{G}}$ und \mathfrak{G} abhängen. Die Darstellung durch diese Matrizen ist jedoch sehr umständlich. Außerdem sehen wir, daß die Invarianten von \mathfrak{B} andere sind als die von \mathfrak{B} . Dies ist insofern unangenehm, als z. B. in der Elastizitätstheorie endlicher Verzerrungen anzunehmen ist, daß die thermodynamischen Größen wie innere Energie, Entropie usw. Funktionen der Invarianten der Verzerrung sind. Wenn dieselben sich nun bei Koordinatentransformation ändern, so ergeben sich hieraus zusätzliche Schwierigkeiten. Auch ist es dann nicht mehr möglich, die Beanspruchung mit Hilfe der Invarianten unabhängig von der Koordinatenwahl zu beschreiben (vgl. § 4).

Die entsprechenden Betrachtungen gelten natürlich für den Fall, daß \mathfrak{B} oder $\hat{\mathfrak{B}}$ zweifach kovariant ist. Es wird daher zweckmäßig sein, auf den mit den ungemischten Tensoren verbundenen Vorteil der Symmetrie zu verzichten.

β) Fall des gemischten Tensors

In dem Falle, daß \mathfrak{B} kovariant-kontravariant ist, muß nach (2.10) sein: $\mathfrak{B} = \overline{\mathfrak{U}}^{-1} \mathfrak{B} \mathfrak{U}$. Wegen (2.3) genügt dann \mathfrak{B} automatisch dem Superpositionsprinzip mit dem gleichen $f(x)$ wie \mathfrak{B} . Insbesondere ist also das eindeutig bestimmte, normierte $f(x)$ unabhängig von der Koordinatenwahl. Weiterhin hat \mathfrak{B} dieselben Invarianten wie \mathfrak{B} . Jede Funktion von \mathfrak{B} , deren Koeffizienten von den Invarianten von \mathfrak{B} abhängig sind, überträgt sich auf die gleiche Funktion von \mathfrak{B} .

Aus der einfachsten Definition $\mathfrak{B} = \mathfrak{Q}$ erhalten wir jetzt bei beliebigen Koordinaten: $\mathfrak{Q}^* = \overline{\mathfrak{U}}^{-1} \mathfrak{Q} \mathfrak{U}$ oder wegen (3.4) und (3.8):

$$\mathfrak{Q}^* = \frac{1}{2} \overline{\mathfrak{U}}^{-1} \ln (\mathfrak{U}^{-1} \mathfrak{A} \hat{\mathfrak{G}}^{-1} \overline{\mathfrak{A}} \overline{\mathfrak{U}}^{-1}) \mathfrak{U}$$

oder

$$(3.12) \quad \mathfrak{Q}^* = \frac{1}{2} \ln (\mathfrak{G} \mathfrak{A} \hat{\mathfrak{G}}^{-1} \overline{\mathfrak{A}}): \text{„logarithmischer Verzerrungstensor“}.$$

\mathfrak{Q}^* gehorcht dem Superpositionsprinzip mit $f(x) \equiv x$.

Der allgemeinste unseren Postulaten genügende Verzerrungstensor ist dann:

$$(3.13) \quad \mathfrak{B} = g(\mathfrak{Q}^*) = h(\sqrt{\mathfrak{G} \mathfrak{A} \hat{\mathfrak{G}}^{-1} \overline{\mathfrak{A}}}) = k(\mathfrak{G} \mathfrak{A} \hat{\mathfrak{G}}^{-1} \overline{\mathfrak{A}}),$$

wo g , h und k den Bedingungen (3.5^a) unterliegen.

Ganz analog erhält man

$$\hat{\mathfrak{B}} = g(\hat{\mathfrak{Q}}^*) = h(\sqrt{\overline{\mathfrak{A}} \mathfrak{G} \mathfrak{A} \hat{\mathfrak{G}}^{-1}}) = k(\overline{\mathfrak{A}} \mathfrak{G} \mathfrak{A} \hat{\mathfrak{G}}^{-1}), \quad \hat{\mathfrak{Q}}^* = \frac{1}{2} \ln (\overline{\mathfrak{A}} \mathfrak{G} \mathfrak{A} \hat{\mathfrak{G}}^{-1}).$$

Bis auf einen beliebigen Faktor ist die Funktion $f(x)$ des Superpositionsprinzips die Umkehrfunktion von $x = g(y)$.

Ganz entsprechend ist vorzugehen, wenn \mathfrak{B} und $\hat{\mathfrak{B}}$ kontravariant-kovariant sein sollen. Es wird dann

$$(3.12^a) \quad \left\{ \begin{array}{l} \mathfrak{Q}^* = \frac{1}{2} \ln (\mathfrak{A} \hat{\mathfrak{G}}^{-1} \overline{\mathfrak{A}} \mathfrak{G}) \\ \text{und} \\ \hat{\mathfrak{Q}}^* = \frac{1}{2} \ln (\hat{\mathfrak{G}}^{-1} \overline{\mathfrak{A}} \mathfrak{G} \mathfrak{A}) \end{array} \right.$$

Alle anderen Beziehungen bleiben ungeändert.

d) Berechnung der Volumdehnung v

Die mit \mathfrak{A} verbundene Volumdehnung ist $v = |\mathfrak{B}|$; also nach (3.7*):

$$v = |\mathfrak{U}|^{-1} \cdot |\hat{\mathfrak{U}}| \cdot |\mathfrak{A}|.$$

Andererseits ist nach (3.6), (3.12) und (3.12^a):

$$|e^{2L^*}| = |\mathfrak{G}| \cdot |\hat{\mathfrak{G}}|^{-1} \cdot |\mathfrak{A}|^2 = v^2.$$

Also ist nach (2.4):

$$(3.14^a) \quad \ln v = \{\mathfrak{L}^*\} = \{\hat{\mathfrak{L}}^*\}$$

oder nach (3.13)

$$(3.14^b) \quad \ln v = \{f(\mathfrak{B})\} = \{f(\hat{\mathfrak{B}})\}.$$

e) Beziehung zum gewöhnlichen Verzerrungstensor

Die gewöhnliche Definition des Verzerrungstensors \mathfrak{T} , resp. $\hat{\mathfrak{T}}$, ist bekanntlich ⁷⁾:

$$ds^2 - d\hat{s}^2 = 2 \cdot d\mathfrak{x} \cdot \mathfrak{T} d\mathfrak{x} = 2 d\hat{\mathfrak{x}} \cdot \hat{\mathfrak{T}} d\hat{\mathfrak{x}}.$$

Nun erhalten wir mit Hilfe von (2.9):

$$ds^2 = d\mathfrak{x} \cdot \mathfrak{G} d\mathfrak{x} = \mathfrak{A} d\hat{\mathfrak{x}} \cdot \mathfrak{G} \mathfrak{A} d\hat{\mathfrak{x}} = d\hat{\mathfrak{x}} \cdot \overline{\mathfrak{A}} \mathfrak{G} \mathfrak{A} d\hat{\mathfrak{x}},$$

und entsprechend

$$d\hat{s}^2 = d\hat{\mathfrak{x}} \cdot \overline{\mathfrak{U}}^{-1} \hat{\mathfrak{G}} \mathfrak{U}^{-1} d\mathfrak{x} = d\hat{\mathfrak{x}} \cdot \hat{\mathfrak{G}} d\hat{\mathfrak{x}}.$$

Also wird:

$$2 \mathfrak{T} = \mathfrak{G} - (\mathfrak{A} \hat{\mathfrak{G}}^{-1} \overline{\mathfrak{A}})^{-1} \quad \text{und} \quad 2 \hat{\mathfrak{T}} = \overline{\mathfrak{A}} \mathfrak{G} \mathfrak{A} - \hat{\mathfrak{G}}.$$

Um die Varianz zu sehen, formen wir nach (3.8) um:

$$(3.15) \quad 2 \mathfrak{T} = \overline{\mathfrak{U}}^{-1} \cdot (\mathfrak{G} - \overline{\mathfrak{U}} \mathfrak{U}^{-1} \hat{\mathfrak{G}} \mathfrak{U}^{-1} \mathfrak{U}) \cdot \mathfrak{U}^{-1} = \overline{\mathfrak{U}}^{-1} \cdot (\mathfrak{G} - \mathfrak{G}^{-2}) \cdot \mathfrak{U}^{-1}$$

und entsprechend

$$(3.15^a) \quad 2 \hat{\mathfrak{T}} = \hat{\mathfrak{U}}^{-1} \cdot (\hat{\mathfrak{G}}^2 - \mathfrak{G}) \hat{\mathfrak{U}}^{-1}.$$

Nach (2.10) sind \mathfrak{T} und $\hat{\mathfrak{T}}$ also zweifach kovariante symmetrische Tensoren. Das Superpositionsprinzip ist für sie nicht erfüllt und die Invarianten ändern sich bei Koordinatentransformation.

Dagegen genügen die gemischten Tensoren $\mathfrak{T} \mathfrak{G}^{-1}$, $\mathfrak{G}^{-1} \mathfrak{T}$, $\hat{\mathfrak{G}}^{-1} \hat{\mathfrak{T}}$ und $\hat{\mathfrak{T}} \hat{\mathfrak{G}}^{-1}$ allen gestellten

Postulaten V 1 bis V 4. Für das Superpositionsprinzip ist nach (3.15) $f(x) = -\frac{1}{2} \ln(1 - 2x)$

für $\mathfrak{T} \mathfrak{G}^{-1}$ und $\mathfrak{G}^{-1} \mathfrak{T}$, dagegen $f(x) = \frac{1}{2} \ln(1 + 2x)$ für $\hat{\mathfrak{G}}^{-1} \hat{\mathfrak{T}}$ und $\hat{\mathfrak{T}} \hat{\mathfrak{G}}^{-1}$ zu setzen.

Für die Volumdehnung v ist dann nach (3.14):

$$v^2 = |\mathfrak{G} - 2 \mathfrak{T} \mathfrak{G}^{-1}|^{-1} = |\mathfrak{G} - 2 \mathfrak{G}^{-1} \mathfrak{T}|^{-1} = |\mathfrak{G} + 2 \hat{\mathfrak{G}}^{-1} \hat{\mathfrak{T}}| = |\mathfrak{G} + 2 \hat{\mathfrak{T}} \hat{\mathfrak{G}}^{-1}|.$$

§ 4. Der Verzerrungsdeviator

a) Postulate

Der Verzerrungsdeviator \mathfrak{D} soll aus dem Verzerrungstensor abgeleitet werden und lediglich die mit der Verzerrung verbundene Gestaltänderung charakterisieren. Damit ergeben sich bereits die an ihn zu stellenden Postulate:

D1: Unterscheiden sich zwei Deformationen nur durch eine Ähnlichkeitsstreckung, so haben sie das gleiche \mathfrak{D} .

D2: Ist mit der Deformation keine Volumänderung verbunden, so ist $\mathfrak{D} = \mathfrak{B}$.

b) Die Realisierung der Postulate

Eine Ähnlichkeitsstreckung im undeformierten oder im deformierten Zustand hat die Gestalt $\lambda \mathfrak{G}$, $\lambda > 0$, mit der Volumdehnung λ^3 . Setzen wir

$$\mathfrak{A} = v^{1/3} \mathfrak{G} \cdot v^{-1/3} \mathfrak{A} = v^{1/3} \mathfrak{G} \cdot \mathfrak{A}_1,$$

so ist nach D1: $\mathfrak{D}(\mathfrak{A}) = \mathfrak{D}(\mathfrak{A}_1)$. Da \mathfrak{A}_1 mit keiner Volumdehnung verbunden ist, ist also:

$\mathfrak{D}(\mathfrak{A}) = \mathfrak{B}(\mathfrak{A}_1) = g(\mathfrak{L}_1^*)$, wobei nach (3.12) und (3.12^a) gilt: $\mathfrak{L}_1^* = -\frac{1}{3} \ln v \cdot \mathfrak{G} + \mathfrak{L}^*$. Hier-

⁷⁾ Vgl. z. B. Moufang l. c.

aus ergibt sich schließlich mit Hilfe von (3.13) und (3.14^a)

$$\mathfrak{D} = g(\mathfrak{Q}^* - \frac{1}{3} \{\mathfrak{Q}^*\} \cdot \mathfrak{E}) = g(f(\mathfrak{B}) - \frac{1}{3} \{f(\mathfrak{B})\} \cdot \mathfrak{E}).$$

Bezeichnen wir allgemein mit $\widetilde{\mathfrak{D}}$ den gewöhnlichen Deviator der Matrix \mathfrak{D} :

$$(4.1) \quad \widetilde{\mathfrak{D}} = \mathfrak{D} - \frac{1}{3} \{\mathfrak{D}\} \cdot \mathfrak{E},$$

so können wir schreiben

$$(4.2) \quad \mathfrak{D} = g(\widetilde{\mathfrak{Q}}^*) = g(\widetilde{f(\mathfrak{B})}).$$

Ganz entsprechend ist

$$\widehat{\mathfrak{D}} = g(\widehat{\mathfrak{Q}}^*) = g(\widehat{f(\mathfrak{B})}).$$

Umgekehrt sind bei dieser Definition offenbar auch die Postulate *D 1* und *D 2* erfüllt. Wird nämlich \mathfrak{A} mit $\lambda > 0$ multipliziert, so geht \mathfrak{Q}^* in $\mathfrak{Q}^* + \ln \lambda \cdot \mathfrak{E}$ über; $\widetilde{\mathfrak{Q}}^*$ und damit \mathfrak{D} bleiben also ungeändert. Ist weiter $v = 0$, so ist nach (3.14^a) $\{\mathfrak{Q}^*\} = 0$, also $\mathfrak{Q}^* = \widetilde{\mathfrak{Q}}^*$ und damit $\mathfrak{D} = g(\mathfrak{Q}^*) = \mathfrak{B}$. Man beachte übrigens, daß \mathfrak{D} automatisch ein Tensor ist, wenn wir \mathfrak{B} als gemischten Tensor verwenden. Auch aus dieser Tatsache ergibt sich eine Bevorzugung der gemischten Tensoren vor den ungemischten.

Am einfachsten wird die Deviatorbildung bei $\mathfrak{B} = \mathfrak{Q}^*$, wo einfach $\mathfrak{D} = \widetilde{\mathfrak{Q}}^*$ wird. Die Verwendung der logarithmischen Deformationsmatrix gestattet also auch bei beliebigen Koordinaten die gewöhnliche Deviatorbildung.

Es sei noch bemerkt, daß bei infinitesimalen Verzerrungen in kartesischen Koordinaten der neue Deviatorbegriff in den alten übergeht. Ist nämlich $\mathfrak{B} = \mathfrak{E} + d\mathfrak{B}$ eine infinitesimale Verzerrung, so ist $\mathfrak{Q} = \frac{1}{2} (d\mathfrak{B} + \overline{d\mathfrak{B}}) + o(d\mathfrak{B})$, so daß (4.2) wegen (3.5^a) liefert:

$$\mathfrak{D} = \widetilde{\mathfrak{Q}} + o(\widetilde{\mathfrak{Q}}) = \frac{1}{2} (d\mathfrak{B} + \overline{d\mathfrak{B}}) + o(d\mathfrak{B}).$$

Für den gewöhnlichen gemischten Verzerrungstensor $\mathfrak{X} \mathfrak{G}^{-1}$ fanden wir in § 3, e: $f(x) = -\frac{1}{2} \ln(1 - 2x)$, also $g(y) = \frac{1}{2} (1 - e^{-2y})$ und $v = |\mathfrak{E} - 2\mathfrak{X} \mathfrak{G}^{-1}|^{-1/2}$. Es ist $\mathfrak{Q}^* = -\frac{1}{2} \ln(\mathfrak{E} - 2\mathfrak{X} \mathfrak{G}^{-1})$, $2\widetilde{\mathfrak{Q}}^* = -\ln(\mathfrak{E} - 2\mathfrak{X} \mathfrak{G}^{-1}) - \frac{2}{3} \ln v \cdot \mathfrak{E}$.

Also wird schließlich ^{a)}:

$$(4.3) \quad \mathfrak{D} = |\mathfrak{E} - 2\mathfrak{X} \mathfrak{G}^{-1}|^{-1/3} \cdot \left(\mathfrak{X} \mathfrak{G}^{-1} - \frac{1}{2} \cdot \left[1 - \sqrt[3]{|\mathfrak{E} - 2\mathfrak{X} \mathfrak{G}^{-1}|} \right] \cdot \mathfrak{E} \right).$$

Entsprechend ist

$$(4.3^a) \quad \widehat{\mathfrak{D}} = |\mathfrak{E} + 2\widehat{\mathfrak{X}} \widehat{\mathfrak{G}}^{-1}|^{-1/3} \left(\widehat{\mathfrak{X}} \widehat{\mathfrak{G}}^{-1} - \frac{1}{2} \cdot \left[\sqrt[3]{|\mathfrak{E} + 2\widehat{\mathfrak{X}} \widehat{\mathfrak{G}}^{-1}|} - 1 \right] \cdot \mathfrak{E} \right).$$

c) Die Verzerrungsinvarianten

Will man den Verzerrungszustand durch Invarianten charakterisieren, so wird man als erste geeignete Invariante von \mathfrak{B} die Volumdehnung oder eine Funktion derselben wählen, während die beiden anderen Invarianten die Gestaltänderung charakterisieren, also bei zusätzlicher Ähnlichkeitsstreckung ungeändert bleiben sollen. Da bei Benutzung der gemischten Tensoren — dies sei im folgenden stets vorausgesetzt — alle Invarianten von \mathfrak{B} auch solche von \mathfrak{Q}^* sind, können wir nach (3.14^a) als erste Invariante $\{\mathfrak{Q}^*\}$ nehmen. Die beiden anderen Invarianten müssen nach Abschnitt (b) dann Invarianten von $\widetilde{\mathfrak{Q}}^*$ sein. Damit ergibt sich die Charakterisierung des Verzerrungszustandes durch

$$(4.4) \quad \begin{cases} j = \{\mathfrak{Q}^*\} & \text{für die Volumänderung} \\ y = \{\widetilde{\mathfrak{Q}}^*\}, z = \{\widetilde{\mathfrak{Q}}^3\} & \text{für die Gestaltänderung.} \end{cases}$$

Wegen $\mathfrak{Q}^* = \overline{\mathfrak{U}}^{-1} \mathfrak{Q} \overline{\mathfrak{U}}$ ist auch $y = \{\widetilde{\mathfrak{Q}}^2\}$ und $z = \{\widetilde{\mathfrak{Q}}^3\}$, so daß y und z die Gestaltänderung unabhängig von der Koordinatenwahl angeben.

^{a)} Man vergleiche die etwas abweichende Bildung bei Moufang, l. c.

Wegen $\{\tilde{\mathfrak{L}}\} = 0$ heißt die charakteristische Gleichung von $\tilde{\mathfrak{L}}$ gemäß (2.2):

$$(4.5) \quad x^3 - \frac{y}{2}x - \frac{z}{3} = 0.$$

Damit diese Gleichung drei reelle Wurzeln hat, muß für

$$(4.6) \quad \zeta = \frac{z^2}{y^3}$$

gelten:

$$(4.7) \quad 0 \leq \zeta \leq \frac{1}{6}.$$

Die geometrische Bedeutung von ζ ergibt sich aus der folgenden Betrachtung. Es sei \mathfrak{S} eine beliebige reine Streckung. Dann können wir \mathfrak{S} auffassen als die n -malige Anwendung der reinen Streckung $\mathfrak{S}_n = \sqrt[n]{\mathfrak{S}}$. Dabei ist $\mathfrak{L}_n = \ln \sqrt[n]{\mathfrak{S}} = \frac{1}{n} \cdot \ln \mathfrak{S} = \frac{1}{n} \mathfrak{L}$; also $\tilde{\mathfrak{L}}_n = \frac{1}{n} \tilde{\mathfrak{L}}$ und damit $y_n = \frac{1}{n^2} \cdot y$ und $z_n = \frac{1}{n^3} z$. Hieraus ergibt sich $\zeta_n = \frac{z_n^2}{y_n^3} = \zeta$. — Ist umgekehrt für zwei Streckungen \mathfrak{S}_1 und \mathfrak{S}_2 : $\zeta_1 = \zeta_2$, so ist $y_1 = \lambda^2 y$ und $z_1 = \lambda^3 z$. Gemäß (4.5) sind dann die Eigenwerte von \mathfrak{S}_1 die λ -te Potenz der Eigenwerte von \mathfrak{S}_2 . Von einer Drehung abgesehen, ist damit $\mathfrak{S}_1 = \mathfrak{S}_2^\lambda$. Wir können uns also \mathfrak{S}_1 und \mathfrak{S}_2 bis auf eine Änderung der Hauptachsen als aus der gleichen infinitesimalen Streckung entstanden denken; bei negativem λ mit Benutzung der Inversen. Dies bedeutet, daß ζ die Beanspruchungsart charakterisiert.

Speziell wird die einachsige volumtreue Verzerrung in geeignet gedrehten kartesischen Koordinaten dargestellt durch

$$\mathfrak{S} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda^{-1/2} & 0 \\ 0 & 0 & \lambda^{-1/2} \end{pmatrix}.$$

Es ist dann
$$\mathfrak{L} = \tilde{\mathfrak{L}} = \ln \lambda \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -1/2 \end{pmatrix};$$

also $y = \ln^2 \lambda \cdot \frac{3}{2}$ und $z = \ln^3 \lambda \cdot \frac{3}{4}$, womit sich ergibt: $\zeta = \frac{1}{6}$.

Dagegen wird für eine reine volumtreue Gleitung

$$\mathfrak{B} = \begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad \text{also} \quad \mathfrak{S}^2 = \mathfrak{B} \mathfrak{B} = \begin{pmatrix} 1 + \lambda^2 & \lambda & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Für die Eigenwerte von \mathfrak{S}^2 erhalten wir aus der charakteristischen Gleichung: $\lambda_1 = 1, \lambda_2 \cdot \lambda_3 = 1$. Für die Eigenwerte von \mathfrak{L} gilt also: $\mu_1 = 0, \mu_2 + \mu_3 = 0$. Damit wird $y > 0, z = 0$; also $\zeta = 0$. Wir haben somit gefunden:

$\zeta = \frac{z^2}{y^3}$ charakterisiert die Beanspruchungsart. Der eine Extremwert $\zeta = 0$ gehört zur reinen Gleitung und der andere Extremwert $\zeta = \frac{1}{6}$ zur einachsigen Streckung.

Die Größe der Beanspruchung wird bei infinitesimalen Verzerrungen üblicherweise durch $\sqrt{\{\mathfrak{D}^2\}}$ charakterisiert. Nun sahen wir, daß für infinitesimale Verzerrungen $\mathfrak{D} \approx \tilde{\mathfrak{L}}$ ist; also gilt \sqrt{y} als Maß für die Größe der Beanspruchung bei infinitesimalen Verzerrungen.

Ist nun wieder \mathfrak{S} eine endliche Streckung, so ist wie oben gezeigt: $\sqrt{y} = n \cdot \sqrt{y_n}$. Da für genügend großes n y_n die Größe der Beanspruchung bei Anwendung der infinitesimalen Verzerrung $\sqrt[n]{\mathfrak{S}}$ angibt, ist es vernünftig, $\sqrt{y} = n \sqrt{y_n}$ als Maß der Beanspruchungsgröße bei n -maliger Anwendung von $\sqrt[n]{\mathfrak{S}}$, also für \mathfrak{S} , zu verwenden. Damit haben wir:

\sqrt{y} gibt die Größe der Beanspruchung an.

§ 5. Der Spannungstensor

Der Spannungstensor \mathfrak{P} soll den Spannungszustand im Punkte \mathfrak{x} des deformierten Materials beschreiben derart, daß bei passender Definition eines Flächenelementes $d\mathfrak{f}$ in \mathfrak{x} die auf

$d\mathbf{f}$ wirkende Kraft durch $\mathfrak{P} d\mathbf{f}$ gegeben wird. Wenn man die Komponenten von \mathfrak{P} mit Hilfe der Transformationsformeln natürlich auch in den Koordinaten von $\hat{\mathbf{e}}$ ausdrücken kann (Übergang zu Lagrangeschen Koordinaten), so bleibt \mathfrak{P} doch an $d\mathbf{f}$ gebunden. Der Versuch, die Spannungen direkt auf $d\mathbf{f}$ zu beziehen, also ein \mathfrak{P} zu konstruieren, ist physikalisch unnatürlich. Wir wollen daher darauf verzichten.

a) Postulate

Bei kartesischen Koordinaten liefert die Spannungsmatrix \mathfrak{P}_0 die auf ein Flächenelement $d\mathbf{f}_0$ im Punkte η wirkende Kraft $d\mathbf{f}_0$ in der Gestalt: $d\mathbf{f}_0 = \mathfrak{P}_0 d\mathbf{f}_0$. Im allgemeinen wird man annehmen können, daß auf das Material keine solchen äußeren Kräfte wirken, die volumabhängige Drehmomente hervorrufen. Es ist bekannt, daß dann \mathfrak{P}_0 symmetrisch ist. Im folgenden braucht jedoch diese Symmetrie nicht vorausgesetzt zu werden.

Haben wir beliebige krummlinige Koordinaten, so muß zunächst ein Flächenelement $d\mathbf{f}$ passend als Transformierte von $d\mathbf{f}_0$ definiert werden. Der Spannungstensor \mathfrak{P} soll dann so gebildet werden, daß $\mathfrak{P} d\mathbf{f}$ wieder die Kraft angibt, die auf das Flächenelement ausgeübt wird. Wird das Element um den Vektor $d\mathbf{z}$ parallel verschoben, so soll die Arbeit $d\mathbf{z} \cdot \mathfrak{P} d\mathbf{f}$ geleistet werden. Damit kommen wir zu den folgenden Postulaten:

P1: \mathfrak{P} ist ein Tensor oder eine tensorielle Dichte.

P2: Für das Flächenelement gilt $d\mathbf{f} = \mathfrak{C} d\mathbf{f}_0$, wo \mathfrak{C} noch passend zu bestimmen ist.

P3: Wird das Flächenelement um $d\mathbf{z}$ verschoben, so wird die Arbeit $dA = d\mathbf{z} \cdot \mathfrak{P} d\mathbf{f}$ geleistet.

b) Die Realisierung der Postulate

dA muß als numerische Größe invariant gegen Koordinatentransformation sein. Also gilt:

$$(5.1) \quad d\mathbf{z} \cdot \mathfrak{P} d\mathbf{f} = d\mathbf{z}_0 \cdot \mathfrak{P}_0 d\mathbf{f}_0,$$

wenn $d\mathbf{z}_0 = \mathfrak{U}^{-1} d\mathbf{z}$ der entsprechende Verschiebungsvektor in kartesischen Koordinaten ist. Nun erhalten wir mit Hilfe von P2 und (2.9):

$$d\mathbf{z}_0 \cdot \mathfrak{P}_0 d\mathbf{f}_0 = \mathfrak{U}^{-1} d\mathbf{z} \cdot \mathfrak{P}_0 \mathfrak{C}^{-1} d\mathbf{f} = d\mathbf{f} \cdot \mathfrak{C}^{-1} \mathfrak{P}_0 \mathfrak{U}^{-1} d\mathbf{z} = d\mathbf{z} \cdot \mathfrak{U}^{-1} \mathfrak{P}_0 \mathfrak{C}^{-1} d\mathbf{f}.$$

Der Vergleich mit (5.1) liefert, da $d\mathbf{z}$ und $d\mathbf{f}$ beliebige Vektoren sind:

$$(5.2) \quad \mathfrak{P} = \mathfrak{U}^{-1} \mathfrak{P}_0 \mathfrak{C}^{-1}.$$

(2.10) zeigt, daß \mathfrak{P} entweder α) zweifach kovariant bei $\mathfrak{C} = \mathfrak{U} \cdot \sqrt{|\mathfrak{G}|}^n$ oder β) kovariant-kontravariant bei $\mathfrak{C} = \mathfrak{U}^{-1} \cdot \sqrt{|\mathfrak{G}|}^n$ ist.

Im Falle α) ist $d\mathbf{f}$ kontravariant und im Falle β) kovariant. Soll die Länge von $d\mathbf{f}$ die geometrische Größe des Flächenelementes sein, so ist $n=0$ zu setzen. \mathfrak{P} ist dann ein echter Tensor; und zwar ist in den Fällen α) und β):

$$(5.2^\alpha) \quad \mathfrak{P} = \mathfrak{U}^{-1} \mathfrak{P}_0 \mathfrak{U}^{-1}$$

und

$$(5.2^\beta) \quad \mathfrak{P} = \mathfrak{U}^{-1} \mathfrak{P}_0 \mathfrak{U}.$$

Im Falle α) ist dann $d\mathbf{f} = \mathfrak{U} d\mathbf{f}_0$. Wird das Flächenelement $d\mathbf{f}_0$ aus den Vektoren $d\mathbf{x}_{10}$ und $d\mathbf{x}_{20}$ gebildet, so ist $d\mathbf{f}_0 = d\mathbf{x}_{10} \times d\mathbf{x}_{20} = \mathfrak{U}^{-1} d\mathbf{x}_1 \times \mathfrak{U}^{-1} d\mathbf{x}_2$ oder mit (2.11): $d\mathbf{f}_0 = \sqrt{|\mathfrak{G}|} \cdot \mathfrak{U} (d\mathbf{x}_1 \times d\mathbf{x}_2)$. Damit wird

$$(5.3^\alpha) \quad d\mathbf{f} = \sqrt{|\mathfrak{G}|} \cdot \mathfrak{G}^{-1} (d\mathbf{x}_1 \times d\mathbf{x}_2).$$

Im Falle β) dagegen ist

$$(5.3^\beta) \quad d\mathbf{f} = \sqrt{|\mathfrak{G}|} \cdot (d\mathbf{x}_1 \times d\mathbf{x}_2).$$

Ob man mit kontravariantem oder kovariantem $d\mathbf{f}$ rechnen will, ist natürlich gleichgültig. Wie (5.3) zeigt, liefert jedoch die kovariante Definition β) die einfachere Formel. Allerdings ist dann \mathfrak{P} nicht mehr symmetrisch, wenn \mathfrak{P}_0 es war. Dafür bleiben jedoch alle Invarianten von \mathfrak{P} bei Koordinatentransformation ungeändert.

c) Die Arbeitsleistung bei infinitesimaler Verzerrung

Wir nehmen jetzt an, es herrsche in einem Raumgebiete um η der homogene durch \mathfrak{P}_0 definierte Spannungszustand. Ein geschlossenes Volumen V habe die Randfläche F mit den Flächenelementen $d\mathbf{f}_0$. Wir führen nun in der Umgebung von η eine homogene infinitesimale

Verzerrung $\mathfrak{E} + d\mathfrak{B}_0$ durch. Das Flächenelement $d\mathfrak{f}_0$ wird dabei einer Verschiebung um den Vektor $d\mathfrak{B}_0 \mathbf{r}_0$ unterworfen, wenn \mathbf{r}_0 sein ursprünglicher Abstand vom Nullpunkt war. Da die gleichzeitige infinitesimale Drehung und Verzerrung des Flächenelementes aus Symmetriegründen keine Arbeitsleistung erfordert, ist die gesamte an dem Volumen geleistete Arbeit

$$V \cdot dA = \iiint_V d\mathfrak{B}_0 \mathbf{r}_0 \cdot \mathfrak{P}_0 d\mathfrak{f}_0 = \iiint_V \bar{\mathfrak{P}}_0 d\mathfrak{B}_0 \mathbf{r}_0 \cdot d\mathfrak{f}_0 = \iiint_V \operatorname{div} (\bar{\mathfrak{P}}_0 d\mathfrak{B}_0 \mathbf{r}_0) dV = \iiint_V \{\bar{\mathfrak{P}}_0 d\mathfrak{B}_0\} dV.$$

Also wird die Arbeit pro Volumeneinheit:

$$(5.4) \quad dA = \{\mathfrak{P}_0 d\mathfrak{B}_0\}.$$

Ist $\mathfrak{E} + d\mathfrak{B}_0$ eine infinitesimale Radialstreckung, also $d\mathfrak{B}_0 = d\lambda \cdot \mathfrak{E}$ mit der Volumendeckung $\frac{dV}{V} = 3 \cdot d\lambda$, so wird $dA = d\lambda \cdot \{\bar{\mathfrak{P}}_0\} = \frac{dV}{V} \cdot \frac{1}{3} \{\bar{\mathfrak{P}}_0\}$. Herrscht die hydrostatische Spannung σ , so wird $dA = \frac{dV}{V} \cdot \sigma$. Auch bei unsymmetrischem \mathfrak{P}_0 vertritt also $\frac{1}{3} \{\bar{\mathfrak{P}}_0\}$ die hydrostatische Spannung, worin die Begründung liegt, $\frac{1}{3} \{\bar{\mathfrak{P}}_0\}$ als mittlere Spannung σ zu bezeichnen.

Für beliebige Koordinaten errechnet sich nach (5.2) dann der Wert der mittleren Spannung zu

$$\sigma = \frac{1}{3} \{\bar{\mathfrak{P}}_0\} = \frac{1}{3} \{\mathfrak{E} \bar{\mathfrak{P}} \mathbb{1}\} = \frac{1}{3} \{\mathbb{1} \mathfrak{E} \bar{\mathfrak{P}}\},$$

also in den Fällen (α) und (β):

$$(5.5^a) \quad \sigma = \frac{1}{3} \{\mathfrak{G}^{-1} \bar{\mathfrak{P}}\} = \frac{1}{3} \{\mathfrak{P} \mathfrak{G}^{-1}\}$$

und

$$(5.5^b) \quad \sigma = \frac{1}{3} \{\bar{\mathfrak{P}}\} = \frac{1}{3} \{\mathfrak{P}\}.$$

Wieder ergibt die gemischt-variante Definition (β) die einfachere Formel.

Der infinitesimalen Verzerrung $\mathfrak{E} + d\mathfrak{B}_0$ entspricht in beliebigen Koordinaten die Verzerrung $\mathfrak{E} + d\mathfrak{B}$ gemäß: $\mathbb{1} (\mathfrak{E} + d\mathfrak{B}_0) d\mathfrak{v}_0 = (\mathfrak{E} + d\mathfrak{B}) \mathbb{1} d\mathfrak{v}_0$, also $d\mathfrak{B}_0 = \mathbb{1}^{-1} d\mathfrak{B} \mathbb{1}$. Wir haben daher mit (5.2) aus (5.4):

$$dA = \{\bar{\mathfrak{E}} \bar{\mathfrak{P}} \mathbb{1} \mathbb{1}^{-1} d\mathfrak{B} \mathbb{1}\} = \{\mathbb{1} \bar{\mathfrak{E}} \bar{\mathfrak{P}} d\mathfrak{B}\}.$$

In den Fällen (α) und (β) wird somit

$$(5.4^a) \quad dA = \{\mathfrak{G}^{-1} \bar{\mathfrak{P}} d\mathfrak{B}\} = \{d\mathfrak{B} \cdot \mathfrak{P} \mathfrak{G}^{-1}\},$$

und

$$(5.4^b) \quad dA = \{\bar{\mathfrak{P}} d\mathfrak{B}\} = \{d\mathfrak{B} \cdot \mathfrak{P}\}.$$

Wieder erhalten wir bei der Definition (β) das einfachere Resultat.

d) Invarianz des Elastizitätsgesetzes

Für isotropes Material hat in kartesischen Koordinaten das Elastizitätsgesetz die Gestalt ⁹⁾:

$$e^i \mathfrak{P}_0 = \frac{\partial e}{\partial j} \mathfrak{E} + 2 \frac{\partial e}{\partial k} \mathfrak{Q} + 3 \frac{\partial e}{\partial l} \mathfrak{Q}^2 \quad \text{bei} \quad j = \{\mathfrak{Q}\}, \quad k = \{\mathfrak{Q}^2\} \quad \text{und} \quad l = \{\mathfrak{Q}^3\},$$

wo e das elastische Potential pro Volumeneinheit des Ausgangszustandes ist.

Wollen wir erreichen, daß diese einfache Gestalt auch für beliebige Koordinaten gilt, so müssen \mathfrak{P} und \mathfrak{Q} die gleiche gemischte Invarianz haben; denn nur dann übertragen sich die Invarianten und funktionellen Abhängigkeiten. Hieraus und aus den weiter oben genannten Gründen erscheint es als am zweckmäßigsten, sowohl \mathfrak{P} als auch \mathfrak{Q} kovariant-kontravariant zu definieren. Aus diesem Grunde ist auch in § 3 diese Varianz bei der Definition von \mathfrak{P} betont worden.

⁹⁾ H. Richter, Das isotrope Elastizitätsgesetz. Z. angew. Math. Mech., Bd. 28 (1948), S. 205—209.

Zur numerischen Auflösung linearer Gleichungssysteme nach dem Matrizenverfahren von Banachiewicz

Herrn Prof. Dr. A. Walther zum 50. Geburtstag

Von Rudolf Zurmühl in Darmstadt

Es wird ein bisher wenig bekanntes allgemein anwendbares Verfahren zur Auflösung linearer Gleichungssysteme beschrieben, welches durch Kombination des Matrizenkalküls mit den in der Rechenmaschine gelegenen Möglichkeiten einer Rechnungsabkürzung zu einer wesentlichen Herabsetzung des Arbeits- und Schreibaufwandes gegenüber den sonst üblichen Methoden führt. Zum Verfahren werden ein Rechenschema mit eingebauten Summenproben sowie bestimmte Regeln der praktischen Maschinenrechnung angegeben und durch ein Zahlenbeispiel erläutert. Das Verfahren eignet sich gleicherweise zur Berechnung der Koeffizientendeterminante und der Kehrmatrix, und es vereinfacht sich wesentlich bei symmetrischem Koeffizientenschema.

A method for the solution of systems of linear equations is described that is till now widely unknown and generally applicable. Combining the calculus of matrices and the possibilities of shortening the work of reckoning based on the use of a calculating machine one succeeds in lessening essentially the computing and writing work, compared with the usual methods. A reckoning scheme with added sum proofs and certain rules for the practical use of the calculating machine are given, and illustrated by a numerical example. The method is apt also for the computation of both the determinant of the coefficients and the reciprocal matrix. It is much simplified, if the matrix of the coefficients is symmetrical.

L'auteur décrit une méthode pour la solution d'un système d'équations linéaires jusqu'à maintenant presque inconnue et d'application générale. En combinant le calcul de matrix et l'usage d'une machine à calculer, on arrive à gagner beaucoup de temps. Un schéma et certaines règles sont donnés et illustrés par un exemple numérique. La méthode peut servir pour le calcul et des déterminants des coefficients et de la matrix réciproque. La matrix des coefficients étant symétrique, le procédé se simplifie.

В этой работе описывается до сих пор мало известный, всегда пригодный метод решения систем линейных уравнений. Благодаря комбинации теории матриц и возможностей уменьшения вычислительной работы при использовании арифмометра этот метод приводит, по сравнению с другими методами, к значительному сокращению вычислительной работы и записи промежуточных результатов. Применительно к этому методу даются расчетная схема с контролем сумм и некоторые правила практического вычисления при помощи арифмометра и все это иллюстрируется на численных примерах. Этот метод пригоден также для вычисления определителей систем уравнений и обратных матриц. Он значительно упрощается в случае симметричной схемы коэффициентов.

1. Einleitung

Die mathematisch durchaus elementare, praktisch jedoch fast immer sehr mühsame Aufgabe der Auflösung umfangreicher linearer Gleichungssysteme hat bekanntlich durch Gauß in Gestalt des Gaußschen Algorithmus eine Lösung gefunden, welche bis heute mit Recht als der vollendete Lösungsweg linearer Gleichungssysteme angesehen worden ist, sofern nicht, was hier außer Betracht bleiben soll, überwiegende Hauptdiagonalglieder der Koeffizientenmatrix eine iterative Behandlung des Gleichungssystems nahelegen. Trotz der hier erreichten weitgehenden Schematisierung aller Rechengänge und der durch systematischen Einbau laufender Kontrollen erzielten Sicherheit zählt die Auflösung umfangreicher Gleichungssysteme indessen auch heute noch zu den gefürchteten Rechenproblemen. Es hat daher nicht an zahlreichen zum Teil höchst geistvollen Vorschlägen gefehlt, den Arbeitsaufwand zu vermindern. Ein kritischer Vergleich dieser Vorschläge mit dem Gaußschen Algorithmus zeigt jedoch immer wieder, daß das Gaußsche Vorgehen allen anderen hinsichtlich der Anzahl der erforderlichen Rechenoperationen überlegen ist¹⁾. Bei der bekannten, von Gauß in seinen Veröffentlichungen geübten Zurückhaltung kann man vermuten, daß Gauß selbst bereits viele dieser Versuche gekannt, ihnen aber den von ihm angegebenen Weg vorgezogen hat. So ist auch schwerlich zu erwarten, daß die Anzahl der zur Auflösung erforderlichen Operationen noch nennenswert herabgesetzt werden kann. Ein Fortschritt in der Lösung der Aufgabe ist daher wohl nur in anderer Richtung zu suchen.

Im Jahre 1938 ist von dem polnischen Astronomen Th. Banachiewicz ein Verfahren veröffentlicht worden²⁾, welches eine entscheidende Herabsetzung des Arbeitsaufwandes

¹⁾ Vgl. etwa O. Heck, Über den Zeitaufwand für das Berechnen von Determinanten und für das Auflösen von linearen Gleichungen. Diss. T. H. Darmstadt 1946. Vgl. aber auch insbesondere den mir erst während der Drucklegung zur Kenntnis gelangten zusammenfassenden Bericht: E. Bodewig, Bericht über die verschiedenen Methoden zur Lösung eines Systems linearer Gleichungen mit reellen Koeffizienten. Proc. Akad. Wetensch. Amsterdam 50 (1947), S. 930–941, 1104–1116, 1285–1295; 51 (1948), S. 53–64, 211–219. Die dort angeführte und gegenüber Gauß ungünstig beurteilte Methode von Boltz-Banachiewicz zur stufenweisen Berechnung einer Kehrmatrix hat mit dem hier zu beschreibenden Verfahren nichts zu tun.

²⁾ Th. Banachiewicz, Méthode de résolution numérique des équations linéaires, du calcul des déterminants et des inverses, et de réduction des formes quadratiques. Bull. intern. Acad. Polon. Sci., Sér. A, 1938, S. 393–404.

überführt, kürzer:

$$\mathfrak{B} \bar{x} = \bar{b} \quad \dots \quad (4')$$

bzw.

$$\overline{\mathfrak{B}} \bar{x} = 0 \quad \dots \quad (4'')$$

mit den Matrizen

$$\mathfrak{B} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ 0 & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_{nn} \end{pmatrix}, \quad \bar{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad \overline{\mathfrak{B}} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} & b_1 \\ 0 & b_{22} & \dots & b_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & b_{nn} & b_n \end{pmatrix} \quad \dots \quad (5).$$

Die Koeffizientenmatrix \mathfrak{A} des Gleichungssystems (1) wird also in eine sogenannte Dreiecksmatrix \mathfrak{B} überführt, deren Elemente unterhalb der Hauptdiagonalen aus Nullen bestehen. Aus dem gestaffelten System (4) können dann die Unbekannten x_i der Reihe nach berechnet werden, anfangend mit x_n aus der letzten, x_{n-1} aus der vorletzten Gleichung usw.

Der Grundgedanke des Verfahrens von Banachiewicz besteht nun darin, die Matrix \mathfrak{A} anzusetzen als das Produkt zweier spezieller Matrizen \mathfrak{B} und \mathfrak{C} in der Form

$$\mathfrak{A} = \mathfrak{C} \mathfrak{B} \quad \dots \quad (6)$$

bzw.

$$\overline{\mathfrak{A}} = \mathfrak{C} \overline{\mathfrak{B}} \quad \dots \quad (6'),$$

wobei \mathfrak{B} bzw. $\overline{\mathfrak{B}}$ im wesentlichen die in (5) angegebene Form der Dreiecksmatrix haben soll, während die Matrix \mathfrak{C} derart gewählt wird, daß die Elemente b_{ik} und c_{ik} schrittweise einzeln aus der Forderung (6) ermittelt werden können. Die Matrix \mathfrak{B} wird noch dahingehend vereinfacht, daß ihre Hauptdiagonalelemente $b_{ii} = 1$ gesetzt werden, wodurch sich die Aufrechnung der Unbekannten aus dem gestaffelten System (4) erleichtert, indem die sonst notwendigen Divisionen durch die b_{ii} fortfallen. Die Matrix \mathfrak{C} , welche nur die Rolle einer Hilfsmatrix zur Bestimmung der allein interessierenden Elemente b_{ik} ($i < k$) spielt, wird nun gleichfalls in Form einer Dreiecksmatrix angesetzt, wobei jetzt die Elemente oberhalb der Hauptdiagonalen aus Nullen bestehen. Die Zerlegung (6') lautet damit ausführlich:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} c_{11} & 0 & 0 & \dots & 0 \\ c_{21} & c_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & c_{n3} & \dots & c_{nn} \end{pmatrix} \begin{pmatrix} 1 & b_{12} & b_{13} & \dots & b_{1n} \\ 0 & 1 & b_{23} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 & b_n \end{pmatrix} \quad \dots \quad (6'').$$

Nach einem bekannten Satz der Determinantenlehre gilt mit (6) für die Determinanten A , B , C der Matrizen \mathfrak{A} , \mathfrak{B} , \mathfrak{C} die Beziehung $A = CB$ oder, da die Determinante von \mathfrak{B} offensichtlich $B = 1$ ist:

$$\boxed{A = C = c_{11} c_{22} \dots c_{nn}} \quad \dots \quad (7).$$

Man erhält also, gleichsam als Nebenprodukt, aus den Diagonalelementen der Hilfsmatrix \mathfrak{C} den Wert der Koeffizientendeterminante A .

Da die Koeffizientenmatrix \mathfrak{A} als nichtsingulär vorausgesetzt wird ($A \neq 0$), so läßt sich das Gleichungssystem (1) stets so anordnen, daß die Matrix \mathfrak{C} die gewünschte Form erhält. Erforderlichenfalls kann man dies durch Umordnen der Spalten von \mathfrak{A} (Umnummerieren der Unbekannten x_i) herbeiführen, vgl. hierzu Nr. 8. Einstweilen sei angenommen, daß die Matrix \mathfrak{C} von selbst die Dreiecksform mit $c_{ii} \neq 0$ annimmt, ohne daß hierzu eine Umordnung der Ausgangsmatrix notwendig ist.

Daß die abgeänderte Matrix $\overline{\mathfrak{B}}$ das Gleichungssystem mit den gesuchten Unbekannten \bar{x}_i erfüllt, folgt aus (1'') mit (6'). Danach ist

$$\overline{\mathfrak{A}} \bar{x} = \mathfrak{C} \overline{\mathfrak{B}} \bar{x} = 0,$$

und da die Matrix \mathfrak{C} mit \mathfrak{A} nichtsingulär ist ($C \neq 0$), so folgt hieraus:

$$\overline{\mathfrak{A}} \bar{x} = 0 \quad \text{bzw.} \quad \mathfrak{A} \bar{x} = \bar{b} \quad \dots \quad (8).$$

3. Bestimmung der Elemente b_{ik} und c_{ik}

Der besondere Bau der Matrizen \mathfrak{B} und \mathfrak{C} erlaubt eine schrittweise Bestimmung der Elemente b_{ik} und c_{ik} aus der Forderung (6''). Die Multiplikation zweier Matrizen \mathfrak{C} und $\overline{\mathfrak{B}}$ in der Reihenfolge $\mathfrak{C}\overline{\mathfrak{B}}$ erfolgt bekanntlich so, daß das Element a_{ik} der Produktmatrix $\overline{\mathfrak{A}} = \mathfrak{C}\overline{\mathfrak{B}}$ gleich dem skalaren Produkt der i -ten Zeile des ersten Faktors \mathfrak{C} mit der k -ten Spalte des zweiten Faktors $\overline{\mathfrak{B}}$ gesetzt wird, wobei die Reihenfolge der Faktoren im allgemeinen nicht

vertauscht werden darf. Danach folgt aus (6'') für das Element a_{ik} der gegebenen Matrix die Beziehung:

$$a_{ik} = c_{i1} b_{1k} + c_{i2} b_{2k} + \dots + c_{il} b_{lk} \quad (l \leq i, k) \dots \dots \dots (9).$$

Dabei ist hier mit Rücksicht auf die Dreiecksform der beiden Faktoren der Index l des letzten von Null verschiedenen Produktes gleich der kleineren der beiden Zahlen i oder k bzw. gleich i und k bei $i = k$.

Man rechnet nun (mit Rücksicht auf eine möglicherweise erforderliche Spaltenumordnung der Ausgangsmatrix) zweckmäßig spaltenweise, beginnend mit der ersten Spalte von \mathfrak{A} und \mathfrak{B} , und durchläuft jede Spalte von oben nach unten. Dabei tritt dann in jeder der aufeinander folgenden Gln. (9) jeweils ein neues, also unbekanntes Element b_{ik} oder c_{ik} auf, und zwar in der k -ten Spalte von \mathfrak{B} und \mathfrak{C} zunächst die Elemente b_{ik} für $i = 1, 2, \dots, k-1$, danach die Elemente c_{ik} für $i = k, k+1, \dots, n$. Auflösen der Gl. (9) nach der jeweiligen Unbekannten b_{ik} (für $i < k$) oder c_{ik} (für $i \geq k$) ergibt die Rechenvorschrift:

$$b_{ik} = (a_{ik} - c_{i1} b_{1k} - c_{i2} b_{2k} - \dots - c_{i,i-1} b_{i-1,k}) : c_{ii} \quad i < k \dots \dots (10a),$$

$$c_{ik} = a_{ik} - c_{i1} b_{1k} - c_{i2} b_{2k} - \dots - c_{i,k-1} b_{k-1,k} \quad i \geq k \dots \dots (10b).$$

Die Vorschrift lautet: Ein Element b_{ik} auf dem Platz i, k ($i < k$) berechnet sich aus dem entsprechenden Element a_{ik} der gegebenen Matrix, vermindert um das skalare Produkt aus allen oberhalb b_{ik} befindlichen Elementen b_{ek} mit den entsprechenden Elementen c_{ie} der betreffenden Zeile i , das Ganze noch dividiert durch das Diagonalelement c_{ii} dieser Zeile. Ein Element c_{ik} auf dem Platz i, k ($i \geq k$) berechnet sich analog aus dem entsprechenden alten Element a_{ik} , vermindert um das skalare Produkt aus allen vor c_{ik} befindlichen Elementen c_{ie} mit den entsprechenden Elementen b_{ek} der betreffenden Spalte k . Wegen $b_{ii} = 1$ erübrigt sich hier eine Division. Die Vorschrift (10) gilt auch noch für die Spalte der rechten Seiten a_i und b_i , wenn man formal setzt $a_i = a_{i, n+1}$, $b_i = b_{i, n+1}$. Sie gilt auch noch für eine unter \mathfrak{A} und \mathfrak{C} anzubringende Kontrollzeilen dienende Summenzeile $-s_k$ und $-\sigma_k$, wenn man setzt $-s_k = a_{n+1, k}$, $-\sigma_k = c_{n+1, k}$, vgl. den folgenden Abschnitt.

Für die erste Spalte der \mathfrak{C} -Matrix folgt sogleich:

$$c_{i1} = a_{i1} \quad (i = 1, 2, \dots, n+1) \dots \dots \dots (11).$$

Für die erste Zeile der \mathfrak{B} -Matrix gilt entsprechend:

$$b_{1k} = a_{1k} : a_{11} \quad (k = 1, 2, \dots, n+1) \dots \dots \dots (12).$$

4. Rechenschema und Proben

Für die praktische Durchführung der Zahlenrechnung vereinigen wir die beiden Dreiecksmatrizen \mathfrak{B} und \mathfrak{C} zu einem einzigen rechteckigen Zahlenschema, wobei in der Hauptdiagonalen nur die Elemente c_{ii} eingetragen werden, während die nicht weiter benötigten Elemente $b_{ii} = 1$ fortgelassen werden können. Dieses Schema, das wir fortan kurz die „ \mathfrak{BC} -Matrix“ nennen wollen, schreiben wir etwa unter das Schema der gegebenen Koeffizientenmatrix \mathfrak{A} und erhalten

Tafel 1. Rechenschema zur Auflösung linearer Gleichungssysteme nach dem Matrizenverfahren

	k-1	2	3	...	n	a_i	Probe
$i = 1$	a_{11}	a_{12}	a_{13}	...	a_{1n}	a_1	?
2	a_{21}	a_{22}	a_{23}	...	a_{2n}	a_2	?
3	a_{31}	a_{32}	a_{33}	...	a_{3n}	a_3	?
...
n	a_{n1}	a_{n2}	a_{n3}	...	a_{nn}	a_n	?
$-s_k$	$-s_1$	$-s_2$	$-s_3$...	$-s_n$	$-s$?
$i = 1$	c_{11}	b_{12}	b_{13}	...	b_{1n}	b_1	x_1
2	c_{21}	c_{22}	b_{23}	...	b_{2n}	b_2	x_2
3	c_{31}	c_{32}	c_{33}	...	b_{3n}	b_3	x_3
...
n	c_{n1}	c_{n2}	c_{n3}	...	c_{nn}	b_n	x_n
$-\sigma_k$	$-\sigma_1$	$-\sigma_2$	$-\sigma_3$...	$-\sigma_n$	$-\sigma = 0$	-1
Probe	?	?	?	...	?	—	—
	$k = 1$	2	3	...	n	b_i	x_i

damit das fertige Rechenschema, Tafel 1, welches, nachdem wir noch die sogleich zu besprechenden Probezeilen und -spalte sowie die Ergebnisspalte der x_i anfügen, alle bei der Rechnung überhaupt benötigten Zahlen bereits in sich vereinigt. Die als Divisoren zu verwendenden Hauptdiagonalglieder c_{ii} heben wir durch Umrahmung hervor.

Für die Brauchbarkeit eines für umfangreiche Zahlenrechnungen bestimmten Verfahrens ist die Möglichkeit ausreichender laufender Rechenkontrollen oft ausschlaggebend. So bilden bekanntlich die Summenproben einen wesentlichen Bestandteil des Gaußschen Algorithmus. Auch beim Matrizenverfahren ist eine solche laufende Kontrolle durch Summenproben möglich⁶⁾. Zu diesem Zweck bildet man etwa die Spaltensummen s_k der Koeffizientenmatrix \mathfrak{A} , also

$$s_k = a_{1k} + a_{2k} + \dots + a_{nk} \dots \dots \dots (13a).$$

Unterwirft man nun diese Elemente s_k dem gleichen Prozeß wie die übrigen Elemente a_{ik} , d. h. denkt man sich die Matrix \mathfrak{A} einfach um eine weitere Zeile mit den s_k erweitert, so erhält man auch in der Hilfsmatrix \mathfrak{C} eine weitere Zeile mit den entsprechenden Elementen σ_k . Diese Größen müssen dann, wie man sich leicht überlegt, gleich den Spaltensummen der \mathfrak{C} -Matrix sein, also

$$\sigma_k = c_{kk} + c_{k+1,k} + \dots + c_{nk} \dots \dots \dots (13b).$$

Praktisch verfährt man besser, ähnlich wie bei Gauß, so, daß man die negativen Werte $-\sigma_k$ und $-\sigma_k$ benutzt und zur Kontrolle die Gesamtsummen

$$c_{kk} + c_{k+1,k} + \dots + c_{nk} - \sigma_k = 0 \dots \dots \dots (14)$$

bildet, welche Null ergeben müssen. Infolge unvermeidlicher Abrundungsfehler können davon abweichend wenige Einheiten der letzten Dezimale auftreten, die man, in Einheiten dieser Dezimalen, in einer letzten Kontrollzeile laufend aufschreibt. Im Rechenschema der Tafel 1 ist zu beachten, daß die Spaltensummierung dabei nur über die \mathfrak{C} -Matrix, also die Glieder auf und unter der Hauptdiagonalen zu erstrecken ist. Dem gleichen Kontrollvorgang werden übrigens auch noch die Spalten der rechten Seiten unterworfen. Das der a_i -Summe s entsprechende Element σ muß dabei Null ergeben, da hier keine c -Elemente mehr auftreten. In Tafel 1 sind die Kontrollplätze durch Fragezeichen angedeutet.

Sind sämtliche Elemente der \mathfrak{BC} -Matrix bestimmt, so beginnt die Aufrechnung der Unbekannten x_i aus dem gestaffelten Gleichungssystem (4) mit $b_{ii} = 1$. Man erhält, anfangend mit der letzten Unbekannten x_n , der Reihe nach aufsteigend:

$$x_{i1} = b_i - x_n b_{in} - x_{n-1} b_{i,n-1} - \dots - x_{i+1} b_{i,i+1} \quad i = n, n-1, \dots, 1 \dots \dots (15).$$

Sind die $n-i$ letzten Unbekannten bereits bestimmt, so errechnet sich die i -te x_i demnach aus der rechten Seite b_i , vermindert um das skalare Produkt aus allen schon vorhandenen Unbekannten mit den Elementen b_{ik} der i -ten Zeile. Man durchläuft die Spalte der x_k und die Zeile der b_{ik} gleichsam in umgekehrter Richtung wie sonst und hört dabei wieder, wie stets, vor dem Diagonalglied der \mathfrak{BC} -Matrix auf.

Sind schließlich alle Unbekannten berechnet, so folgt noch die Probe durch Einsetzen der n Unbekannten x_i in das Ausgangsgleichungssystem gemäß

$$a_{i1} x_1 + a_{i2} x_2 + \dots + a_{in} x_n - a_i = 0 \quad i = 1, 2, \dots, n+1 \dots \dots \dots (16),$$

wobei $i = n+1$ sich auf die letzte Summenzeile der \mathfrak{A} -Matrix bezieht. Man bildet also wiederum skalare Produkte aller Zeilen der \mathfrak{A} -Matrix mit der um die Komponente -1 verlängerten Ergebnisspalte der x_i . Die Ergebnisse der Proben, also Nullen bzw. wenige Einheiten der letzten Dezimalen, trägt man in die der \mathfrak{A} -Matrix angefügten Probepalte ein. Bei ungünstiger Koeffizientenmatrix (relativ sehr kleine Koeffizientendeterminante) können die Einsetzungsreste übrigens gelegentlich auch beträchtliche Werte annehmen, ohne daß die Rechnung falsch sein müßte; doch ist dies ein seltener Ausnahmefall.

5. Die Zahlenrechnung

Zählt man die beim Matrizenverfahren erforderlichen Einzeloperationen, also die Anzahl der Additionen, der Multiplikationen und Divisionen ab, so kommt man ungefähr auf die gleiche Anzahl wie beim Gaußschen Algorithmus. Wollte man also etwa die Multiplikationen einzeln vornehmen, die Ergebnisse aufschreiben und alles addieren, so hätte das Verfahren gegenüber dem bewährten Gaußschen Vorgehen nicht den geringsten Vorteil. Eine Arbeitersparnis ergibt sich erst dann, wenn man die für das Verfahren charakteristischen skalaren Produkte unmittelbar mit einer Rechenmaschine, d. h. durch Auflaufen-

⁶⁾ Die Summenproben finden sich in der unter ²⁾ zitierten Hauptarbeit noch nicht, wohl aber in anderem Zusammenhang in einer weiteren Arbeit: T. H. Banachiewicz, La règle de Chio, cracoviens et matrices. Bull. intern. Acad. Polon. Sci. Sér. A, 1938, S. 405—412.

lassen aller Teilprodukte zum Gesamtwert im Ergebniswerk der Maschine ausführt, ohne Ablesen und Niederschreiben der Einzelprodukte. Nur in dieser Verbindung des Matrizenkalküls mit der durch die Eigenart der Rechenmaschine bedingten Rechnungsabkürzung kann die entscheidende Überlegenheit des Matrizenverfahrens gegenüber den bisherigen Rechenmethoden gesehen werden. Das Verfahren steht und fällt mit der sachgemäßen Benutzung der Rechenmaschine. Demgemäß ist die Beachtung einiger einfacher Regeln und Kunstgriffe für ein rasches und sicheres Arbeiten wesentlich, die indessen leicht zu erlernen sind, so daß auch der im Maschinenrechnen Ungeübte sich bald hinreichend einarbeitet.

Zunächst ist es mit Rücksicht auf einheitliche Kommastellung ratsam, die rechten Seiten a_i , falls notwendig, durch Multiplizieren mit einer geeigneten Zehnerpotenz auf die gleiche Größenordnung wie die übrigen Koeffizienten a_{ik} zu bringen gemäß $\bar{a}_i = 10^m a_i$. Das bedeutet den Ersatz der Unbekannten x_i durch neue Werte $\bar{x}_i = 10^m x_i$, welche von der Größenordnung der 1 sind. Alle Elemente der \bar{A} -Matrix erhalten dann die gleiche Stellenzahl nach dem Komma.

Wie die Gln. (10) und insbesondere (11) und (12) erkennen lassen, sind die Elemente c_{ik} von der Größenordnung der Koeffizienten a_{ik} , die Elemente b_{ik} dagegen von der Größenordnung der 1. Die beiden Zahlenarten der \mathcal{BC} -Matrix werden also im allgemeinen nicht mit gleicher Stellenzahl nach dem Komma, sondern, mit Rücksicht auf volle Ausnutzung der Genauigkeit, mit gleicher Anzahl geltender Stellen gerechnet, z. B. durchweg mit 6 oder 7 Stellen, vgl. das Zahlenbeispiel in Nr. 6.

Mit Rücksicht auf die bei den b_{ik} erforderliche Division empfiehlt es sich, bei der Ausrechnung der skalaren Produkte in (10a) und (10b) die Faktoren c_{ik} grundsätzlich im Einstellwerk (Tastatur, Einstellhebel) der Maschine einzustellen, die Faktoren b_{ik} hingegen dann wie üblich in das Umdrehungszählwerk laufen zu lassen. Das erste Glied a_{ik} wird wie ein c -Element behandelt, multipliziert mit dem Faktor 1. Die Ergebnisse c_{ik} werden dann unmittelbar im Ergebniswerk der Maschine (im Maschinenwagen) abgelesen. Für die Ergebnisse b_{ik} ist der gleichfalls im Ergebniswerk erscheinende Klammerwert der Gl. (10a) noch durch das Diagonalglied c_{ii} zu dividieren, was durch Eintasten des c_{ii} -Wertes wie bei den anderen c -Werten und anschließende Ausführung der Division geschieht. Das Ergebnis b_{ik} erscheint dann in der üblichen Weise im Umdrehungszählwerk der Maschine. Bei Beachtung dieser Regeln arbeitet man mit fester Kommastellung bei der gesamten Berechnung der \mathcal{BC} -Matrix.

Die Produkte werden mit den in den Formeln (10) angegebenen unter Beachtung der den Zahlen eigenen Vorzeichen gerechnet, wobei sich negative Produkte selbsttätig abziehen. Negative Ergebnisse bei den c_{ik} erscheinen dann als Komplementzahlen, z. B. ...9943,8752 = -56,1248. Wird der Klammerwert bei b_{ik} negativ, so wird die Division nicht wie üblich durch Abziehen, sondern durch Hinzulassen des Divisors c_{ii} bis auf die Null ausgeführt bei positiv (nicht wie sonst bei der Division negativ) zählendem Umdrehungszählwerk, wobei freilich eine automatische Division nicht mehr, wohl aber noch „Stoppdivision“ möglich ist. Das Vorzeichen von b_{ik} richtet sich dann natürlich noch nach dem von c_{ii} .

Die Ergebnisse \bar{x}_i sind von der gleichen Größenordnung wie die b_{ik} . Bei der Bildung der skalaren Produkte in (15) müssen also die Kommaeinstellungen auf der Maschine u. U. abgeändert werden. Die Probe nach (16) erfolgt dann wieder mit der alten Stellenmarkierung, wobei die a_{ik} in das Einstellwerk, die \bar{x}_i in das Umdrehungszählwerk gegeben werden.

An Rechenmaschinen eignet sich grundsätzlich jede der üblichen nach dem Additionsprinzip arbeitenden Maschinen. Doch ist bei Maschinen, welche nicht über eine durchgehende Zehnerübertragung im Ergebniswerk verfügen (Staffelwalzensystem wie Rheinmetall-Borsig oder Archimedes), Vorsicht geboten, da hier das Ergebnis beim Auflaufen der skalaren Produkte durch die auf der Vorderseite des Maschinenwagens aussetzende Zehnerübertragung leicht verfälscht werden kann, was indessen bei einiger Aufmerksamkeit im gegebenen Augenblick bemerkt wird. Alle übrigen Maschinenbauarten (Sprossenradsystem wie Brunswiga, Zahnstangensystem wie Mercedes-Euklid, Schaltklinkensystem wie Hamann) sind ohne weiteres und mit völliger Sicherheit verwendbar.

6. Ein Zahlenbeispiel

Gegeben sei das lineare Gleichungssystem für 4 Unbekannte:

$$\begin{aligned} 43,57x_1 - 23,84x_2 - 51,65x_3 + 19,44x_4 &= 0,7362 \\ 62,89x_1 + 84,97x_2 + 21,84x_3 - 39,35x_4 &= 1,1872 \\ 37,48x_1 + 93,24x_2 + 84,39x_3 + 26,75x_4 &= 0,3875 \\ -19,37x_1 + 54,38x_2 + 14,59x_3 + 62,85x_4 &= 0,5738 \end{aligned}$$

Die rechten Seiten werden durch Multiplikation mit 100 auf die gleiche Größenordnung wie die Koeffizienten gebracht, man rechnet also mit abgeänderten Unbekannten $\bar{x}_i = 100x_i$. Nach

Anhängen zweier Schutzstellen wird die Rechnung durchweg sechs- bis siebenstellig durchgeführt. In Zahlentafel 1 ist die gesamte Zahlenrechnung enthalten. Die Zahl für c_{33} in der dritten Zeile und dritten Spalte entsteht beispielsweise aus

$$c_{33} = 84,3900 - 37,4800 \cdot -1,185449 - 113,7477 \cdot 0,807438 = 36,9764.$$

Die Zahl b_3 in der dritten Zeile der b_i -Spalte entsteht aus

$$b_3 = (38,7500 - 37,4800 \cdot 1,689695 - 113,7477 \cdot 0,104330) : 36,9764 = -0,985685.$$

Zahlentafel 1. Zahlenbeispiel zur Auflösung linearer Gleichungssysteme nach dem Matrizenverfahren

	$k=1$	2	3	4	q_i	Probe $\cdot 10^4$
$i=1$	43,5700	-23,8400	-51,6500	19,4400	73,6200	0
2	62,8900	84,9700	21,8400	-39,3500	118,7200	-1
3	37,4800	93,2400	84,3900	26,7500	38,7500	0
4	-19,3700	54,3800	14,5900	62,8500	57,3800	0
$-s_k$	-124,5700	-208,7500	-69,1700	-69,6900	-288,4700	+1
$i=1$	43,5700	0,547165	-1,185449	0,446179	1,689695	0,637671
2	62,8900	119,3812	0,807438	-0,564663	0,104330	1,404457
3	37,4800	113,7477	36,9764	2,008209	-0,985685	-1,448886
4	-19,3700	43,7814	-43,7229	184,0189	0,230654	0,230654
$-\sigma_k$	-124,5700	-276,9003	6,7465	-184,0189	0,0000	-1
Probe	0	0	0	0		
	$k=1$	2	3	4	b_i	x_i

Die Ergebnisse sind, auf vier Stellen abgerundet, entsprechend den vierstellig gegebenen Ausgangswerten:

$$\begin{aligned} x_1 &= 0,006377 \\ x_2 &= 0,014045 \\ x_3 &= -0,014489 \\ x_4 &= 0,002307 \end{aligned}$$

7. Vereinfachungen bei symmetrischer Matrix

Wie bei den meisten übrigen Verfahren zur Gleichungslösung ergeben sich auch beim Matrizenverfahren Vereinfachungen im Falle symmetrischer Koeffizientenmatrix γ). Dabei ist nämlich:

$$b_{ki} = \frac{c_{ik}}{c_{kk}} \dots \dots \dots (17),$$

d. h. auch die $\mathcal{B}\mathcal{C}$ -Matrix ist in gewisser Hinsicht symmetrisch, wenn man auf der \mathcal{B} -Seite, also oberhalb der Hauptdiagonalen noch die Division durch die Diagonalelemente c_{kk} berücksichtigt. Bei der Berechnung eines Elementes c_{ik} wird demnach das gespiegelte Element b_{ki} der $\mathcal{B}\mathcal{C}$ -Matrix einfach dadurch mitgeliefert, daß man das Ergebnis c_{ik} noch durch das in der gleichen Spalte darüber befindliche Diagonalelement c_{kk} dividiert. Eine besondere Berechnung der b -Elemente erübrigt sich somit mit Ausnahme der rechten Seiten b_i .

Die Richtigkeit dieser Behauptung ergibt sich durch Induktionsschluß wie folgt unter Beachtung der Symmetriebedingung $a_{ik} = a_{ki}$. Für die erste Spalte und Zeile der $\mathcal{B}\mathcal{C}$ -Matrix gilt:

$$\begin{aligned} c_{i1} &= a_{i1} \\ c_{11} b_{1i} &= a_{1i} = a_{i1} = c_{i1}. \end{aligned}$$

Für die zweite Spalte und Zeile liefern die Gl. (10) hiermit:

$$\begin{aligned} c_{i2} &= a_{i2} - c_{i1} b_{12} = a_{i2} - c_{i1} \frac{c_{21}}{c_{11}} \\ c_{22} b_{2i} &= a_{2i} - c_{21} b_{1i} = a_{i2} - c_{21} \frac{c_{i1}}{c_{11}} = c_{i2} \end{aligned}$$

⁷⁾ Hierauf hat Herr Dipl.-Math. Gerh. Schreier, Darmstadt aufmerksam gemacht. Bei Banachiewicz selbst findet sich darüber nichts.

Für die dritte Spalte und Zeile gilt damit:

$$c_{i3} = a_{i3} - c_{i1} b_{13} - c_{i2} b_{23} = a_{i3} - c_{i1} \frac{c_{31}}{c_{11}} - c_{i2} \frac{c_{32}}{c_{22}},$$

$$c_{33} b_{3i} = a_{3i} - c_{31} b_{1i} - c_{32} b_{2i} = a_{3i} - c_{31} \frac{c_{i1}}{c_{11}} - c_{32} \frac{c_{i2}}{c_{22}} = c_{i3} \text{ usf.}$$

8. Abweichungen vom Regelfall

Das Verfahren in der bisherigen Form ohne Zusatzmaßnahmen versagt, sobald eines der Diagonalelemente $c_{ii} = 0$ wird, was genau dann eintritt, wenn eine Hauptabschnittsdeterminante der Ausgangsmatrix \mathfrak{A} verschwindet. Da andererseits $A = C \neq 0$ sein soll, kann der Fall $c_{ii} = 0$ nur in der Weise eintreten, daß die Spalten von \mathfrak{C} gegenüber der normal angenommenen Dreiecksform vertauscht sind. Das aber kann dann durch ein entsprechendes Vertauschen in der Ausgangsmatrix \mathfrak{A} behoben werden. Im Rechenschema, Tafel 1, bedeutet dies lediglich ein Abändern der Nummern k am Kopf des Schemas, wobei jetzt freilich die abgeänderten Nummern mit den entsprechenden unveränderten Nummern k am Fuß des Schemas in Verbindung zu bringen sind. Sobald also im Laufe der Rechnung in einer k -ten Spalte das Diagonalelement $c_{kk} = 0$ wird, vertauscht man die beiden Nummern k und $k + 1$ am Kopf der \mathfrak{A} -Matrix und rechnet jetzt die k -te Spalte der $\mathfrak{B}\mathfrak{C}$ -Matrix mit der $(k + 1)$ -ten (jetzt mit k bezeichneten) Spalte der \mathfrak{A} -Matrix, und ebenso die $(k + 1)$ -te Spalte der $\mathfrak{B}\mathfrak{C}$ -Matrix mit der k -ten (jetzt mit $k + 1$ bezeichneten) Spalte der \mathfrak{A} -Matrix. Dabei müssen natürlich die bereits berechneten b_{ik} -Werte oberhalb des Elementes c_{kk} ausradiert und neu gerechnet werden. Die Umnumerierung muß dann zum Schluß auch bei der Reihenfolge der Unbekannten berücksichtigt werden. Verschwindet nach der Spaltenvertauschung das Diagonalelement noch einmal, so ist die Vertauschung zu wiederholen, d. h. die k -te Spalte der \mathfrak{A} -Matrix ist durch die $(k + 2)$ -te Spalte zu ersetzen usf. Da bei einer Vertauschung zweier Spalten die Determinante einer Matrix ihr Vorzeichen umkehrt, so berechnet sich die Koeffizientendeterminante jetzt zu

$$A = (-1)^z C \dots \dots \dots (18),$$

wenn z die Anzahl der Spaltenvertauschungen angibt.

Die Spaltenvertauschung wird man übrigens auch dann vornehmen, wenn ein Diagonalelement c_{kk} im Vergleich zu den übrigen c -Werten auffallend klein wird. Auf diese Weise schützt man sich vor u. U. beträchtlichen Genauigkeitsverlusten.

Beispiel:

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 &= 13 \\ 4x_1 + 6x_2 - 2x_3 &= -4 \\ -x_1 + 3x_2 + 2x_3 &= 1 \end{aligned}$$

	1	3	2	a_i	Probe
1	2	3	4	13	0
2	4	6	-2	-4	0
3	-1	3	2	1	0
—σ	-5	-12	-4	-10	0
1	2	2	1,5	6,5	2
2	4	-10	0	3	3
3	-1	4	4,5	-1	-1
—σ	-5	6	-4,5	0	-1
Probe	0	0	0	—	—
	1	2	3	b_i	x_i

Ergebnis: $x_1 = 2, x_2 = -1, x_3 = 3.$
 $A = C \cdot (-1)^1 = +90.$

9. Berechnung der Kehrmatrix

Die Berechnung der Kehrmatrix (der inversen Matrix) $\mathfrak{A}^{-1} = (\alpha_{ik})$ zu einer gegebenen Koeffizientenmatrix $\mathfrak{A} = (a_{ik})$, definiert durch

$$\mathfrak{A} \mathfrak{A}^{-1} = \mathfrak{A}^{-1} \mathfrak{A} = \mathfrak{E} \dots \dots \dots (19)$$

mit der Einheitsmatrix

$$\mathfrak{E} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \dots \dots \dots (20),$$

vollzieht sich beim Matrizenverfahren in der gleichen Weise wie beim Gaußschen Algorithmus einfach dadurch, daß man die Auflösung eines Gleichungssystems mit der Matrix \mathfrak{A} und mit den n -Spalten der Einheitsmatrix an Stelle der einen Spalte a_i als rechten Seiten durchführt. Im Rechenschema der Tafel 1 treten also an die Stelle der einen Spalte a_i die n -Spalten der Einheitsmatrix. Entsprechend ergeben sich an Stelle der einen Spalte b_i der $\mathfrak{B}\mathfrak{C}$ -Matrix jetzt n -Spalten. Aus ihnen berechnen sich dann n Ergebnisspalten mit n^2 -Zahlenwerten x_{ik} , deren Schema unmittelbar die gesuchte Kehrmatrix darstellt gemäß

$$\alpha_{ik} = x_{ik} \dots \dots \dots (21).$$

10. Zum Arbeitsbedarf

Will man den Arbeitsbedarf des Matrizenverfahrens vergleichen etwa mit dem des Gaußschen Algorithmus, so kann man zunächst einfach die Anzahl der niederzuschreibenden bzw. der der Rechenmaschine zu entnehmenden Zahlenwerte abzählen. Macht man dabei auch im Gaußschen Algorithmus, soweit möglich, von der Abkürzung durch Auflaufenlassen von Teilprodukten im Ergebniswerk der Rechenmaschine Gebrauch, so kommt man unter Einrechnung aller Summenproben sowie der Aufrechnung der Unbekannten und der Einsetzungsproben zu folgenden Anzahlen z . Einschließlich der gegebenen Koeffizienten a_{ik} und der rechten Seiten a_i sind niederzuschreiben:

beim Matrizenverfahren: $z_M = 2n^2 + 7n + 3 = (2n + 1)(n + 3)$

beim Gauß-Algorithmus: $z_G = \frac{n}{6}(2n^2 + 15n + 19) = \frac{n}{6}(z_M + 8n + 16)$.

Zieht man beide Male die Anzahl $n(n + 1)$ der gegebenen Werte a_{ik} und a_i ab, so erhält man die Anzahlen z' der der Maschine zu entnehmenden Zahlenwerte. In Zahlentafel 2 sind diese Anzahlen für einige Werte der Zahl n der Unbekannten zusammengestellt. Die Anzahlen verringern sich natürlich bei beiden Verfahren, wenn die Koeffizientenmatrix, was praktisch häufig der Fall ist, teilweise aus Nullen besteht.

Zahlentafel 2. Vergleich der Anzahlen der niederzuschreibenden bzw. der der Rechenmaschine zu entnehmenden Zahlenwerte beim Matrizenverfahren (z_M) und beim Gaußschen Algorithmus (z_G)

Zahl der Unbekannten n	Niederzuschreibende Zahlen		Der Rechenmaschine zu entnehmende Ergebnisse	
	z_M	z_G	z'_M	z'_G
3	42	41	30	29
4	63	74	42	54
5	88	120	58	90
10	273	615	163	505
20	943	3 730	523	3 310
50	5 353	48 075	2 803	45 525
100	20 703	358 650	10 603	348 550

Das Verhältnis des tatsächlichen Arbeitsbedarfes ist natürlich keineswegs so kraß, wie es nach dieser eindrucksvollen Zusammenstellung den Anschein haben könnte, da die Berechnung jedes Zahlenwertes beim Matrizenverfahren im Mittel eine wesentlich größere Anzahl von Einzeloperationen verlangt als bei Gauß. Verlässliche Angaben hierüber sind indessen nicht ohne eine sorgfältige Analyse der Elementarzeiten möglich, unterstützt durch umfangreiche praktische Rechenversuche etwa im Sinne der unter 1) zitierten Arbeit von Heck.

Ein nicht zu unterschätzender Vorzug des Matrizenverfahrens besteht übrigens noch in der völligen Gleichförmigkeit aller Rechengänge, nämlich der fortgesetzten und ausschließlichen Bildung skalarer Produkte, wobei sich auch die zur Berechnung der b_{ik} erforderlichen Divisionen durch die Diagonalelemente gleichsam als letztes Glied dieser Produktbildung glatt in den übrigen Rechnungsablauf einfügen, sofern man die oben angeführten Hinweise zur Maschinenrechnung beachtet. Hieraus ergeben sich möglicherweise Vorteile für eine rein maschinelle Auflösung linearer Gleichungssysteme durch Rechenautomaten (Hollerith u. dgl.).

Eingegangen: 15. 5. 48.

Zur Bestimmung reeller Nullstellen von reellen Funktionen einer Variablen

Von Erwin Pflanz in Stuttgart.

Obere und untere Schranken für eine Nullstelle x^* einer Funktion $y = f(x)$ kann man — Monotonie der Ableitung $f^{(N)}(x)$ in der Umgebung von x^* vorausgesetzt — durch die Nullstellen gewisser, $f(x)$ dort „einschließender“ Polynome N -ten Grades in x erhalten. Für dieses Verfahren wird u. a. nachgewiesen, wie sich seine Konvergenz beschleunigen läßt und daß es auch bei Nichtzutreffen der Monotonievoraussetzung nach einfacher Abänderung Schranken für x^* liefert.

Upper and lower limits of a zero x^* of a function $y = f(x)$ can be obtained by the zeroes of certain polynomials of the N -th order in x that are including $f(x)$ in the vicinity of x^* , provided that the derivative $f^{(N)}(x)$ is monotonous. Among other things, it is proved how to accelerate the convergence and that the method furnishes limits of x^* after a slight modification, even if the supposition of the monotony is not realized.

On peut arriver aux limites inférieures et supérieures d'un zéro x^* d'une fonction $y = f(x)$ à l'aide de certains polynômes du n -ième ordre en x , qui enferment $f(x)$ dans le voisinage de x^* , à condition que la dérivée $f^{(N)}(x)$ soit monotone. On montre comment on peut accélérer la convergence et on prouve que la méthode fournit (après une simple modification) des limites pour x^* , même sans supposer une monotonie.

Верхняя и нижняя границы корня x^* функции $y = f(x)$ могут быть получены в случае монотонности производной $f^{(N)}(x)$ в окрестности x^* при помощи корней некоторых полиномов степени N , заключающих в этой области функцию $f(x)$. Кроме того показывается, как может быть ускорена сходимость этого метода и как путем незначительного его изменения могут быть получены границы для x^* также и в случае, если условие монотонности не выполнено.

1. Einleitung. Die Funktion $y = f(x)$ sei in einem Intervall $J(x_1 \leq x \leq x_2)$ eindeutig und besitze dort genau eine Nullstelle x^* . Es sei $f(x_1) \cdot f(x_2) < 0$. Obere und untere Schranken für x^* sind durch die Nullstellen gewisser Interpolationspolynome $y = p_{\alpha_1, \alpha_2}(x)$ gegeben. Dabei sei in J $f^{(n-1)}(x)$ monoton und $f^{(q)}(x)$ ($q = 1, 2, \dots, n$) dort stetig, weiter sei $n = \alpha_1 + \alpha_2 \geq 3$, $\alpha_1 \cdot \alpha_2 \geq 1$ und $y = p_{\alpha_1, \alpha_2}(x)$ dasjenige (eindeutig bestimmte) Polynom höchstens $(n-1)$ ten Grades in x , das mit $f(x)$ in $x = x_\mu$ ($\mu = 1; 2$) im Funktionswert und in den aufeinanderfolgenden Ableitungen nach x bis zur $(\alpha_\mu - 1)$ ten übereinstimmt. ($f^{(0)}(x_\mu)$ bedeute $f(x_\mu)$). Zu diesem, in einem früheren Aufsatz¹⁾ betrachteten Verfahren der Wurzelbestimmung durch Einschließung, einer Verallgemeinerung des kombinierten Verfahrens Newton-Regula falsi, seien hier einige Bemerkungen hinzugefügt.

Für den einfachsten Fall $n = 3$ (bei $\alpha_1 = 2, \alpha_2 = 1$ bzw. $\alpha_1 = 1, \alpha_2 = 2$) zeigt die Darstellung²⁾

$$\left. \begin{aligned} f(x) - p_{2,1}(x) &= -\frac{1}{3!} (x - x_1)^2 (x_2 - x) f'''(z_0), \\ f(x) - p_{1,2}(x) &= \frac{1}{3!} (x - x_1) (x_2 - x)^2 f'''(\bar{z}_0) \end{aligned} \right\} \dots \dots \dots (1);$$

$$(x_1 \leq x \leq x_2, x_1 < z_0, \bar{z}_0 < x_2)$$

daß für alle x in J

$$\left. \begin{aligned} \text{entweder stets } p_{1,2}(x) \leq f(x) \leq p_{2,1}(x) \text{ gilt, falls dort ständig } f'''(x) \geq 0 \\ \text{oder } p_{2,1}(x) \leq f(x) \leq p_{1,2}(x) \text{ „ „ „ „ „ „ „ „ } f'''(x) \leq 0 \text{ ist} \end{aligned} \right\} \dots \dots \dots (2).$$

Für in J monotonen $f''(x)$ schließen also die Parabeln 2.0. $p_{1,2}(x)$ und $p_{2,1}(x)$ dort $f(x)$ ein. Bei $f(x_1) \cdot f(x_2) < 0$ sind somit durch die im Intervallinnern gelegene Nullstelle von $y = p_{2,1}(x)$ bzw. die von $y = p_{1,2}(x)$, die wir mit $x_{2,1}^*$ bzw. $x_{1,2}^*$ bezeichnen wollen, obere und untere Schranken für x^* gegeben.

Der Unterschied dieser Schranken ist (bei $f'(x^*) \neq 0$) klein in h ($2h = x_2 - x_1 > 0$) von dritter Ordnung. Durch einfache Umbildung von $f(x)$ läßt sich, wie in Abschnitt 2 erläutert wird, erreichen, daß die so sich ergebende neue Abweichung $x_{2,1}^* - x_{1,2}^*$ mit der fünften Potenz von h gegen Null geht. Daß auch bei Nichterfülltsein der eben erwähnten Monotoniebedingung für $f''(x)$ das Einschließungsverfahren sich durch Übergang zu einem geeigneten Gleichungssystem noch durchführen läßt, wird in Abschnitt 3 gezeigt. In Abschnitt 4 geben wir Beispiele für die Gewinnung von Schranken für x^* bei Einschließung von $y = f(x)$ durch in y quadratische Polynome. Schließlich werden in Abschnitt 5 bei Zugrundelegung von drei Interpolationsstellen weitere Näherungsformeln hergeleitet. — Die nachstehenden Überlegungen lassen sich auch für $n = 4, 5, \dots$ durchführen. Die dabei sich ergebenden Formeln sind jedoch im allgemeinen für

¹⁾ E. Pflanz, Z. angew. Math. Mech. 28 (1948), S. 114—122.

²⁾ A. Markoff, Differenzenrechnung (deutsche Übersetzung von Th. Friesendorff und E. Prümmer). Leipzig 1896, S. 8.

den praktischen Gebrauch nicht genügend einfach. Wir gehen daher nicht näher darauf ein und beschränken uns auf diesen Hinweis.

2. Konvergenzverbesserung. Es sei $f(x_1) \cdot f(x_2) < 0$, $f''(x)$ monoton in $J(x_1 \leq x \leq x_2)$, $f'''(x)$ dort stetig. x^* sei die einzige Nullstelle von $f(x)$ im Innern von J . Dann gilt³⁾, bei $f'(x) \neq 0$ in J und für kleine $h = \frac{1}{2}(x_2 - x_1)$

$$x_{2,1}^* - x_{1,2}^* = \xi_{2,1}^* - \xi_{1,2}^* = -2 \frac{h^3}{3!} \frac{f_m'''}{f_m'}, \left[1 - \left(\frac{\xi_{1,2}^*}{h} \right)^2 \right] + \dots, \left(\left| \frac{\xi_{1,2}^*}{h} \right| < 1 \right) \dots (3),$$

mit $\xi_{2,1}^* = x_{2,1}^* - x_m, \dots, x_m = \frac{1}{2}(x_1 + x_2)$, $f_m' = f'(x_m), \dots$, wobei die in (3) weggelassenen Glieder von höherer als dritter Ordnung in h sind. Nach (3) wird das x^* -einschließende Intervall $x_{2,1}^* - x_{1,2}^*$ und damit der Fehler der Näherungen für x^* (für kleine h) besonders klein sich ergeben, wenn $f'''(x_m) = f_m''' = 0$ ist; denn dann verschwindet die Abweichung $x_{2,1}^* - x_{1,2}^*$ von höherer als dritter Ordnung. Der Ansatz

$$v(x) = (1 + b_1 \xi + b_2 \xi^2) \cdot f(x), \quad (\xi = x - x_m) \dots (4)$$

gibt nun bei passender Festlegung der Konstanten b_1, b_2 eine Funktion $v(x)$, für welche $v_m''' = v'''(x_m) = 0$ ist und die wie $f(x)$ für $x = x^*$ verschwindet⁴⁾. Neben der Forderung

$$v_m''' = 0 \dots (5)$$

muß, um der erwünschten Monotonie von $v''(x)$ in J nicht zu widersprechen, auch

$$v_m^{(4)} = 0 \dots (6)$$

sein, denn sonst ändert $v'''(x)$ in der Umgebung von $x = x_m$ sein Vorzeichen.

(5) und (6) ergeben für die Konstanten b_1, b_2 des Ansatzes (4)

$$b_1 = \frac{f_m' f_m^{(4)} - 2 f_m'' f_m'''}{2(3 f_m''^2 - 2 f_m' f_m''')}, \quad b_2 = \frac{4 f_m''^3 - 3 f_m'' f_m^{(4)}}{12(3 f_m''^2 - 2 f_m' f_m''')} \dots (7).$$

Weiter erhält man [auf gleiche Weise wie die Entwicklung (3)] für das zu $v(x)$ nach (4) und (7) gehörige, neue Schrankenintervall

$$x_{2,1}^* - x_{1,2}^* = \xi_{2,1}^* - \xi_{1,2}^* = -4 \frac{h^5}{5!} \frac{v_m^{(5)}}{v_m'} \left[1 - \left(\frac{\xi_{1,2}^*}{h} \right)^2 \right] + \dots, \quad (v_m' = v'(x_m) \neq 0) \dots (8).$$

Wegen der Gl. (6) bis (8) sei in J $f(x)$ fünfmal stetig nach x differenzierbar vorausgesetzt.

Bei dem Beispiel $f(x) = e^x - 2$ mit $x_1 = 0, x_2 = 1$ ist $b_1 = -\frac{1}{2}, b_2 = \frac{1}{12}$. Damit ergibt sich $v'''(x) = \frac{1}{12} e^x \cdot \xi^2 \geq 0$, also $v''(x)$ monoton für $-\frac{1}{2} \leq \xi \leq \frac{1}{2}, 0 \leq x \leq 1$ und es ist dort $1 + b_1 \xi + b_2 \xi^2 \neq 0$. Nach (8) ist die Abweichung $x_{2,1}^* - x_{1,2}^* \approx -1,5 \cdot 10^{-4}$ zu erwarten. Die Auflösung der quadratischen Gleichungen $p_{2,1}(x) = 0$ bzw. $p_{1,2}(x) = 0$ führt auf die Ungleichung $x_{2,1}^* = 0,69307 \dots \leq x^* = \ln 2 = 0,69314 \dots \leq 0,69321 \dots = x_{1,2}^*$, gegenüber dem Ergebnis⁵⁾ $0,673 < x^* < 0,703$ bei der gewöhnlichen Einschließung von $f(x)$ (also bei $b_1 = b_2 = 0$) durch Polynome $p_{1,2}(x), p_{2,1}(x)$.

Bei dem weiteren Beispiel $f(x) = x^3 - 3$ ($1 < x^* < 2$) erhält man gleicherweise durch Übergang zu $v(x)$, mit $x_1 = 1, x_2 = 2$ $1,4415 \dots \leq x^* \leq 1,4430 \dots$, und durch Wiederholung des Verfahrens mit $x_1 = 1,44, x_2 = 1,45$ unter Beibehaltung der Werte b_1, b_2 des ersten Schrittes $1,442249570_{18} \dots \leq x^* \leq 1,442249570_{73} \dots$. Mit $b_1 = b_2 = 0$ ergibt sich bei $x_1 = 1, x_2 = 2$ hier nur $1,425 \dots \leq x^* \leq 1,463 \dots$.

Zur Berechnung von b_1, b_2 nach (7) benötigt man die Ableitungen von $f(x)$ bis zur vierten Ordnung in der Intervallmitte. Man kann jedoch anstatt b_1, b_2 auch Ausdrücke \bar{b}_1, \bar{b}_2 verwenden, die nur die ersten und zweiten Ableitungen von $f(x)$, und zwar an den Intervallenden, enthalten. Die Abweichung $\xi_{2,1}^* - \xi_{1,2}^*$ der Polynomnullstellen, welche zu der mit \bar{b}_1, \bar{b}_2 ge-

³⁾ Siehe a. a. O. ¹⁾, Gl. (24a).

⁴⁾ Um die Konvergenz bei der Anwendung des gewöhnlichen Verfahrens von Newton und der erweiterten Newtonschen Formeln zu beschleunigen, wurde von Fr. A. Willers (Z. angew. Math. Mech. 16 (1936), S. 315—316, und 18 (1938), S. 197—200) eine Methode angegeben, mit der unser oben erläutertes Verfahren den Ausgangspunkt gemeinsam hat. Zur Verbesserung der Konvergenz sucht man in der Reihenentwicklung für die ν -te Abweichung d_ν (einer ν -ten Näherung für x^*) nach Potenzen von $d_{\nu-1}$ den Faktor der niedersten Potenz von $d_{\nu-1}$ möglichst klein zu machen bzw. zum Verschwinden zu bringen. Bei Willers geschieht dies durch Transformation der Variablen x , bei unserer Methode durch Umbildung der Funktion.

⁵⁾ a. a. O. ¹⁾, Abschn. 4c).

bildeten Funktion $\bar{v}(x) = (1 + \bar{b}_1 \xi + \bar{b}_2 \xi^2) f(x)$ gehört, geht dann, wie wir zeigen werden, gleichfalls mit der fünften Potenz von h gegen Null. Zum Nachweis entwickelt man $[f'_+ = f'(x_2), f'_- = f'(x_1), \dots, |\partial_\nu| < 1, \nu = 1, \dots, 4]$

$$\left. \begin{aligned} f'_+ &= f'_m + \dots + \frac{h^4}{24} f^{(5)}(x_m + \partial_1 h), f'_- = f'_m - \dots - \frac{h^4}{24} f^{(5)}(x_m + \partial_2 h), \\ f''_+ &= f''_m + \dots + \frac{h^3}{6} f^{(5)}(x_m + \partial_3 h), f''_- = f''_m - \dots - \frac{h^3}{6} f^{(5)}(x_m + \partial_4 h) \end{aligned} \right\} \quad (9).$$

Die Auflösung von (9) nach f'_m, \dots ergibt mit den Abkürzungen

$$f'_+ + f'_- = \sigma', f''_+ + f''_- = \sigma'', f'_+ - f'_- = \delta', f''_+ - f''_- = \delta''$$

die Entwicklungen

$$\left. \begin{aligned} f'_m &= \frac{1}{2} \sigma' - \frac{1}{4} h \delta'' + h^4 \cdot S_1, & h f''_m &= \frac{3}{4} \delta' - \frac{1}{4} h \sigma'' + h^4 \cdot S_2 \\ h^2 f'''_m &= \frac{1}{2} h \delta'' + h^4 \cdot S_3, & h^3 f^{(4)}_m &= -\frac{3}{2} \delta' + \frac{3}{2} h \sigma'' + h^4 \cdot S_4 \end{aligned} \right\} \quad (10).$$

In (10) sind S_1, \dots gewisse, hier näher nicht interessierende lineare Summenausdrücke über die in (9) auftretenden fünften Ableitungen von f . Setzt man f'_m, \dots aus (10), unter Vernachlässigung der S_ν , in (7) ein, so erhält man die gesuchten Ausdrücke

$$\left. \begin{aligned} \bar{b}_1 &= -\frac{1}{h} \frac{[6 \sigma' \delta' - 3 h (2 \sigma' \sigma'' - \delta' \delta'') + h^2 \sigma'' \delta'']}{[27 \delta'^2 - 2 h (9 \delta' \sigma'' + 4 \sigma' \delta'') + h^2 (3 \sigma''^2 + 4 \delta''^2)]}, \\ \bar{b}_2 &= \frac{1}{6 h^2} \frac{[27 \delta'^2 - 36 h \delta' \sigma'' + h^2 (9 \sigma''^2 + 8 \delta''^2)]}{[27 \delta'^2 - 2 h (9 \delta' \sigma'' + 4 \sigma' \delta'') + h^2 (3 \sigma''^2 + 4 \delta''^2)]}. \end{aligned} \right\} \quad (11).$$

Für Polynome $f(x)$ vom Grade ≤ 4 ist (wegen $S_1 = S_2 = \dots = 0$) $\bar{b}_1 = b_1, \bar{b}_2 = b_2$.

Der Zusammenhang zwischen b_1 und \bar{b}_1, b_2 und \bar{b}_2 läßt sich für kleine h darstellen durch

$$\bar{b}_1 = b_1 + h^3 \cdot \varphi_3 + \dots, \bar{b}_2 = b_2 + h^2 \cdot \psi_2 + \dots,$$

wobei die (für kleine h beschränkten) Ausdrücke φ_3, ψ_2 die Funktionswerte f'_+, f'_-, f''_+, f''_- und die S_1, \dots enthalten.

Für die Nullstellen $\bar{x}_{2,1}^*, \bar{x}_{1,2}^*$ bzw. $\bar{\xi}_{2,1}^*, \bar{\xi}_{1,2}^*$ der zu $\bar{v}(x)$ gehörigen Einschließungspolynome $p_{2,1}, p_{1,2}$ folgt dann

$$\bar{\xi}_{2,1}^* - \bar{\xi}_{1,2}^* = \bar{x}_{2,1}^* - \bar{x}_{1,2}^* = -4 \frac{h^5}{5!} \left[\frac{\bar{v}_m^{(5)}}{\bar{v}_m} + \frac{5!}{2} \cdot \frac{f'_m}{\bar{v}_m} \cdot \psi_2 \right] \left[1 - \left(\frac{\bar{\xi}_{1,2}^*}{h} \right)^2 \right] + \dots, (\bar{v}_m \neq 0) \quad (12).$$

Durch Heranziehung der Funktionen $v(x), \bar{v}(x)$ anstatt $f(x)$ und bei Einschließung von $v(x), \bar{v}(x)$ durch Parabeln 2.0. $p_{2,1}(x), p_{1,2}(x)$ erhält man also die gleiche Güte der Annäherung (in h) von x^* durch die Parabelnullstellen $\bar{x}_{2,1}^*, \bar{x}_{1,2}^*$ wie bei Einschließung von $f(x)$ durch gewisse Polynome⁶⁾ vierten Grades in x . Bei Verwendung von $v(x), \bar{v}(x)$ hat man anstatt Gleichungen vierten Grades jedoch nur solche zweiten Grades aufzulösen⁷⁾.

3. Einschließung der Nullstelle x^* von $y = f(x)$ bei Nichterfülltsein der Monotonievoraussetzung. Es sei $f(x)$ in $J(x_1 \leq x \leq x_2)$ eindeutig, $f(x_1) f(x_2) < 0$, x^* die einzige Nullstelle von $f(x)$ in $J, f^{(\nu)}(x)$ ($\nu = 1, 2, 3$) dort stetig. Über die Monotonie von $f''(x)$ sei nichts vorausgesetzt. Die Nullstellen der zugehörigen Polynome $p_{1,2}(x), p_{2,1}(x)$ werden also möglicherweise in J x^* nicht einschließen. Wir werden zeigen, daß sich dann trotzdem mit dem Einschließungsverfahren Schranken für x^* finden lassen und zwar durch Übergang zu einem passend gewählten Gleichungssystem.

Wir können jedenfalls (dreimal stetig differenzierbare) Funktionen $\varphi(x), \psi(x)$ angeben, so daß

$$f(x) = \varphi(x) - \psi(x). \quad (13)$$

ist und in $x_1 \leq x \leq x_2$ $\varphi''(x)$ und $\psi''(x)$ monoton sind.

⁶⁾ a. a. O. 1), Gl. (24b).

⁷⁾ Gleichfalls auf die Auflösung einer quadratischen Gleichung kommt man (siehe R. Beyer, Z. angew. Math. Mech. 16 (1936), S. 345—346), wenn man $f(x)$ durch einen Kegelschnitt annähert, der in einem Nachbarnpunkt x_1 von x^* $f(x)$ fünfpunktig berührt. Der Fehler, der damit für x^* erzielten Näherung ist dann klein in δ von fünfter Ordnung, wenn man mit δ die Abweichung der Ausgangsnäherung x_1 von x^* bezeichnet.

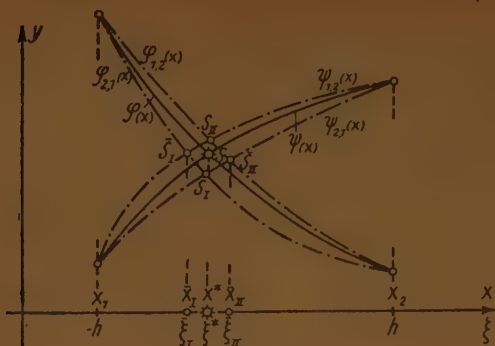


Bild 1

$y = \varphi(x)$ und $y = \psi(x)$ liegt (s. Bild 1) in dem von den vier Parabelbögen begrenzten Viereck $\bar{S}_I, \bar{S}_I, \bar{S}_{II}, \bar{S}_{II}$.

Dabei sei

$$\begin{aligned} \bar{S}_I & \left(\text{mit der Abszisse } \bar{x}_I \text{ bzw. } \bar{\xi}_I = \frac{x_1 + x_2}{2} - \bar{x}_I \right) \text{ der Schnittpunkt von } y = \varphi_{2,1}(x) \text{ und } y = \psi_{1,2}(x) \\ \bar{S}_{II} & \left(\text{,, ,, ,, } \bar{x}_{II} \text{ ,, } \bar{\xi}_{II} = \frac{x_1 + x_2}{2} - \bar{x}_{II} \right) \text{ ,, ,, ,, } y = \varphi_{1,2}(x) \text{ ,, } y = \psi_{2,1}(x) \\ \bar{S}_I & \left(\text{,, ,, ,, } x_I \text{ ,, } \xi_I = \frac{x_1 + x_2}{2} - x_I \right) \text{ ,, ,, ,, } y = \varphi_{2,1}(x) \text{ ,, } y = \psi_{2,1}(x) \\ \bar{S}_{II} & \left(\text{,, ,, ,, } x_{II} \text{ ,, } \xi_{II} = \frac{x_1 + x_2}{2} - x_{II} \right) \text{ ,, ,, ,, } y = \varphi_{1,2}(x) \text{ ,, } y = \psi_{1,2}(x) \end{aligned}$$

x^* wird (s. Bild 1) von mindestens einem der beiden Paare x_1, x_{II} bzw. \bar{x}_I, \bar{x}_{II} eingeschlossen und zwar, falls $f''(x)$ in J nicht monoton ist, jedenfalls von \bar{x}_I, \bar{x}_{II} , wie wir zeigen werden.

Aus (1) folgt, wenn man darin $f(x)$ durch $\varphi(x), \psi(x)$ ersetzt und $p_{1,2}(x)$ bzw. $p_{2,1}(x)$ durch $\varphi_{1,2}(x), \psi_{1,2}(x)$ bzw. $\varphi_{2,1}(x), \psi_{2,1}(x)$,

$$\begin{aligned} \varphi_{2,1}(x) - \psi_{1,2}(x) &= \varphi(x) - \psi(x) + \frac{1}{3!} (x - x_1)(x_2 - x) \{ (x - x_1) \varphi'''(\delta_2) + (x_2 - x) \psi'''(\delta_3) \} \\ \varphi_{1,2}(x) - \psi_{2,1}(x) &= \varphi(x) - \psi(x) - \frac{1}{3!} (x - x_1)(x_2 - x) \{ (x_2 - x) \varphi'''(\delta_1) + (x - x_1) \psi'''(\delta_4) \} \end{aligned} \quad (14a)$$

$$\begin{aligned} \varphi_{2,1}(x) - \psi_{2,1}(x) &= \varphi(x) - \psi(x) + \frac{1}{3!} (x - x_1)^2 (x_2 - x) \{ \varphi'''(\delta_2) - \psi'''(\delta_4) \} \\ \varphi_{1,2}(x) - \psi_{1,2}(x) &= \varphi(x) - \psi(x) - \frac{1}{3!} (x - x_1)(x_2 - x)^2 \{ \varphi'''(\delta_1) - \psi'''(\delta_3) \} \end{aligned} \quad (14b)$$

In (14a), (14b) sind (bei Beschränkung von x auf das Intervall $x_1 \leq x \leq x_2$) $\delta_1, \delta_2, \dots$ gewisse Zwischenstellen zwischen x_1 und x_2 . Wegen der vorausgesetzten Monotonie von $\varphi'''(x), \psi'''(x)$ wechseln in unserem Intervalle $\varphi'''(x), \psi'''(x)$ ihre jeweiligen Vorzeichen nicht. Es besitzen also dort $\varphi'''(x), \psi'''(x)$ entweder (a) durchweg gleiches Vorzeichen oder (b) durchweg entgegengesetztes Vorzeichen. (Das bloße Verschwinden einzelner Funktionswerte werde nicht als Vorzeichenwechsel angesehen.) Wenn $f''(x)$ in J nicht monoton ist, kommt (b) nach (13) nicht in Frage, so daß der Fall (a) zutrifft. Die Gl. (14a) zeigen, daß dann x^* durch \bar{x}_I, \bar{x}_{II} eingeschlossen wird. Ist im Falle (a) überdies $(\varphi'''(z_1) - \psi'''(z_2))$ durchweg positiv oder durchweg negativ für alle z_1, z_2 in $x_1 \leq z_1, z_2 \leq x_2, \dots$, so folgt aus (14b), daß auch x_1, x_{II} obere und untere Schranken für x^* geben. Im Falle (b) schließen (nach (14b)) x_1, x_{II} die gesuchte Nullstelle x^* ein.

(Bei $\varphi'''(x) \equiv 0$ bzw. $\psi'''(x) \equiv 0$ ist $\varphi(x) \equiv \varphi_{1,2}(x) \equiv \varphi_{2,1}(x)$ bzw. $\psi(x) \equiv \psi_{1,2}(x) \equiv \psi_{2,1}(x)$; offensichtlich fallen dann x_I mit x_{II} , \bar{x}_I mit \bar{x}_{II} zusammen.)

x_1, \dots, x_{II} bzw. ξ_1, \dots, ξ_{II} ergeben sich (als Abszissen der Schnittpunkte der Einschließungspolynome $y = \varphi_{2,1}(x), y = \psi_{2,1}(x), \dots$) aus den zwischen $-h < \xi < h$ gelegenen Nullstellen von Polynomen $\Phi_{2,1}(\xi) - \Psi_{2,1}(\xi) = 0, \dots, \Phi_{1,2}(\xi) - \Psi_{1,2}(\xi) = 0$ ($\xi = x - \frac{x_1 + x_2}{2}$), also als Lösungen von quadratischen Gleichungen der Form

$$k_0 + 2k_1(\xi/h) + k_2(\xi/h)^2 = 0 \quad (15).$$

Die zur jeweiligen Wurzel ξ_I, \dots gehörigen Konstanten k_0, k_1, k_2 sind dem nachstehenden Schema zu entnehmen.

	k_0	$k_1 \neq 0$	k_2
(x_{II}, ξ_{II})	$3\delta_+ + \delta_- - 2h\delta'_+$	$\left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \delta_+ - \delta_-$	$-(\delta_+ - \delta_-) + 2h\delta'_+$
(x_I, ξ_I)	$3\delta_- + \delta_+ + 2h\delta'_-$		$\delta_+ - \delta_- - 2h\delta'_-$
$(\bar{x}_{II}, \bar{\xi}_{II})$	$3\delta_{\pm} + \delta_{\mp} - 2h\sigma'_{\pm}$		$-(\delta_{\pm} - \delta_{\mp}) + 2h\sigma'_{\pm}$
$(\bar{x}_I, \bar{\xi}_I)$	$3\delta_{\mp} + \delta_{\pm} + 2h\sigma'_{\mp}$		$\delta_{\pm} - \delta_{\mp} - 2h\sigma'_{\mp}$

wobei

$$\begin{aligned} \delta_+ &= \varphi_+ - \psi_+, \quad \delta'_+ = \varphi'_+ - \psi'_+, \\ \delta_- &= \varphi_- - \psi_-, \quad \delta'_- = \varphi'_- - \psi'_-, \\ \delta_{\pm} &= \varphi_+ - \psi_-, \quad \delta_{\mp} = \varphi_- - \psi_+, \\ \sigma'_{\pm} &= \varphi'_+ + \psi'_-, \quad \sigma'_{\mp} = \varphi'_- + \psi'_+ \end{aligned}$$

und

$$\varphi_- = \varphi(x_I), \quad \varphi_+ = \varphi(x_2),$$

$\varphi'_+ = \varphi'(x_2), \dots$ bedeuten.

Für die durch Auflösung von (15) sich ergebenden, in $-h < \xi < h$ liegenden Wurzeln ξ_I, ξ_{II} bzw. $\bar{\xi}_I, \bar{\xi}_{II}$ erhält man für kleine h

$$\left. \begin{aligned} \xi_I - \xi_{II} &= x_I - x_{II} = -2 \frac{h^3}{3!} \cdot \frac{(\varphi'''_m - \psi'''_m)}{(\varphi'_m - \psi'_m)} \left[1 - \left(\frac{\xi_{II}}{h} \right)^2 \right] + \dots \\ \bar{\xi}_I - \bar{\xi}_{II} &= \bar{x}_I - \bar{x}_{II} = -2 \frac{h^3}{3!} \cdot \frac{(\varphi'''_m + \psi'''_m)}{(\varphi'_m - \psi'_m)} \left[1 - \left(\frac{\xi_{II}}{h} \right)^2 \right] + \dots \end{aligned} \right\} (\varphi'_m \neq \psi'_m) \quad (16).$$

In (16) ist $\varphi'_m = \varphi'(x_m), \dots, x_m = \frac{x_1 + x_2}{2}$.

Die speziell für $\psi \equiv 0$ aus (13) bis (16) folgenden Formeln stimmen mit den früher⁸⁾ (für den Fall einer monotonen Funktion $f''(x)$) angegebenen Ausdrücken überein.

4. Einschließung von $y = f(x)$ durch Parabeln mit der Gl. $x = \sum_{v=0}^2 c_v y^v$. Die in $x_1 \leq x \leq x_2$ eindeutige Funktion $y = f(x)$ besitze dort genau eine Nullstelle x^* , es sei $f(x_1) \cdot f(x_2) < 0$. Weiter sei in $x_1 \leq x \leq x_2$ $f(x)$ dreimal stetig nach x differenzierbar und

$$f'(x) \neq 0 \quad (17).$$

Es bedeute $y_+ = f(x_2) = f_+, y_- = f(x_1) = f_-, y_1 = \text{Min}(y_+, y_-), y_2 = \text{Max}(y_+, y_-)$. Wegen (17) besitzt $y = f(x)$ eine in $y_1 \leq y \leq y_2$ eindeutige Umkehrfunktion $x = g(y)$. Es seien $x = q_{2,1}(y) = \sum_{v=0}^2 c_v^{(2,1)} y^v$ bzw. $x = q_{1,2}(y) = \sum_{v=0}^2 c_v^{(1,2)} y^v$ die Polynome (höchstens) 2ten Grades in y , die durch die Punkte $(x_1, f(x_1)), (x_2, f(x_2))$ gehen und die Umkehrfunktion $x = g(y)$ in $(x_1, f(x_1))$ bzw. $(x_2, f(x_2))$ in 1. Ordnung berühren (s. Bild 2). $q_{2,1}(y), q_{1,2}(y)$ schließen dann in $y_1 \leq y \leq y_2$ die Umkehrfunktion $g(y)$ ein, wenn dort $x = g(y)$ dreimal stetig nach y differenzierbar und

$$\frac{d^2 g}{dy^2} = -\frac{f''(x)}{f'^3(x)} \text{ monoton} \quad (18)$$

ist. Bei unseren Voraussetzungen über $f(x)$ ist die erste Bedingung erfüllt.

Man erhält so x^* einschließende Schranken x_I, x_{II} sofort durch $q_{2,1}(0), q_{1,2}(0)$, zu deren Bestimmung ersichtlich keine quadratischen Gleichungen aufzulösen sind. Andererseits ist, wie aus Stetigkeitsgründen für $h \rightarrow 0$ folgt, nur für kleine h gewährleistet, daß x_I, x_{II} ins Innere des Ausgangsintervalles (x_1, x_2) fallen.

Es ergibt sich

$$\left. \begin{aligned} x_I &= c_0^{(2,1)} = x_1 + 2h \left(\frac{f_-}{f_- - f_+} \right)^2 \left[1 + \left(\frac{f_- - f_+}{f_-} \right) \frac{f_+}{2hf_-} \right] \\ x_{II} &= c_0^{(1,2)} = x_2 - 2h \left(\frac{f_+}{f_+ - f_-} \right)^2 \left[1 - \left(\frac{f_+ - f_-}{f_+} \right) \frac{f_-}{2hf_+} \right] \end{aligned} \right\} \dots (x_2 - x_1 = 2h) \quad (19).$$

Mit $x_2 - x^* = \Theta_2 h, x^* - x_1 = \Theta_1 h, 0 < \Theta_1, \Theta_2 < 2, \Theta_1 + \Theta_2 = 2$ folgt hieraus für kleine h

$$x_I - x_{II} = \Theta_1 \Theta_2 h^3 \left[\frac{f''_2}{f'^2_2} - \frac{1}{3} \frac{f'''}{f'^3} \right] + \dots, \quad (f = f'(x^*), \dots),$$

wobei $0 < \Theta_1 \cdot \Theta_2 \leq 1$ ist.

⁸⁾ a. a. O. 1).

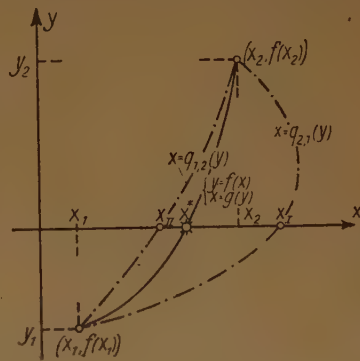


Bild 2

Für das im Abschn. 2 betrachtete Beispiel $f(x) = x^3 - 3$ mit $x_1 = 1,44$, $x_2 = 1,45$ erhält man aus (19)

$$x_{II} = 1,442249_{46} \dots \leq x^* \leq 1,442249_{60} \dots = x_I.$$

Bei Annäherung von $x = g(y)$ durch Taylorsche Polynome (2. Grades in y) $q_{3,0}(y)$ bzw. $q_{0,3}(y)$, welche $g(y)$ in den Punkten $(x_1, f(x_1))$ bzw. $(x_2, f(x_2))$ jeweils dreipunktig berühren, ergeben sich wie es sein muß, die dreigliedrigen erweiterten Newtonschen Formeln

$$x_I = x_1 - \frac{f_-}{f'_-} - \frac{1}{2} \frac{f_-^2 f''_-}{f'^3_-}, \quad x_{II} = x_2 - \frac{f_+}{f'_+} - \frac{1}{2} \frac{f_+^2 f''_+}{f'^3_+} \dots \dots (20).$$

Hinsichtlich Formeln für die Abweichungen $x_I - x^*$, $x_{II} - x^*$ der Näherungen (20) sei auf die Literatur⁹⁾ verwiesen.

Die Paare x_I , x_{II} aus (19), (20) schließen x^* ein, wenn obige Voraussetzungen über $f(x)$, insbesondere (17), (18) erfüllt sind.

5. Näherungsformeln bei 3 Interpolationsstellen $x_1, x_m = \frac{1}{2}(x_1 + x_2), x_2$. Die in $x_1 \leq x \leq x_2$ eindeutige und viermal stetig nach x differenzierbare Funktion $f(x)$ besitze dort genau eine Nullstelle x^* . Es sei $f(x_1)f(x_2) < 0$.

a) $y = p_{2,1,1}(x)$ bzw. $y = p_{1,1,2}(x)$ seien die durch die Forderungen

$$p_{2,1,1}(x_1) = f_-, \quad p'_{2,1,1}(x_1) = f'_-, \quad p_{2,1,1}(x_m) = f(x_m) = f_m, \quad p_{2,1,1}(x_2) = f_+$$

bzw.

$$p_{1,1,2}(x_1) = f_-, \quad p_{1,1,2}(x_m) = f_m, \quad p_{1,1,2}(x_2) = f_+, \quad p'_{1,1,2}(x_2) = f'_+$$

eindeutig festgelegten Polynome 3. Grades in x .

Aus der (bei Markoff¹⁰⁾ angegebenen) Darstellung für das Fehlerglied $f(x) - p_{1,1,2}(x)$ bei drei Interpolationsstellen ersieht man, daß die in $x = x_m$ sich durchsetzenden Parabeln $p_{2,1,1}(x)$, $p_{1,1,2}(x)$ in $x_1 \leq x \leq x_2$ die Kurve $y = f(x)$ einschließen, falls dort $f'''(x)$ monoton ist.

Die Gleichungen der Einschließungspolynome lauten ($\xi = x - x_m$, $2h = x_2 - x_1$)

$$y = P_{2,1,1}(\xi) = \sum_{v=0}^3 \delta_v \left(\frac{\xi}{h}\right)^v, \quad y = P_{1,1,2}(\xi) = \sum_{v=0}^3 \gamma_v \left(\frac{\xi}{h}\right)^v$$

mit

$$\begin{aligned} \delta_0 &= f_m & \gamma_0 &= f_m \\ 4\delta_1 &= f_+ + 4f_m - 5f_- - 2hf'_- & 4\gamma_1 &= -(f_- + 4f_m - 5f_+ + 2hf'_+) \\ 2\delta_2 &= f_+ - 2f_m + f_- & \gamma_2 &= \delta_2 \\ 4\delta_3 &= f_+ - 4f_m + 3f_- + 2hf'_- & 4\gamma_3 &= -(f_- - 4f_m + 3f_+ - 2hf'_+) \end{aligned}$$

Bei hinreichend kleinem h besitzen $P_{2,1,1}(\xi)$ bzw. $P_{1,1,2}(\xi)$ jeweils genau eine Nullstelle $\xi_{2,1,1}^*$ bzw. $\xi_{1,1,2}^*$ in $-h < \xi < h$, welche, falls obige Monotonievoraussetzung hinsichtlich $f'''(x)$ erfüllt ist, x^* einschließen.

Weiter gilt. ($f'_* = f'(x^*) \neq 0$, $x_{2,1,1}^* = \xi_{2,1,1}^* + x_m$, ...)

$$\xi_{2,1,1}^* - \xi_{1,1,2}^* = x_{2,1,1}^* - x_{1,1,2}^* = -2 \frac{h^4}{4!} \frac{f''''_*}{f'^4_*} \left[1 - \left(\frac{\xi_{1,1,2}^*}{h} \right)^2 \right] \frac{\xi_{1,1,2}^*}{h} + \dots$$

wobei (wegen $|\xi_{1,1,2}^*|/h < 1$)

$$\left| \frac{\xi_{1,1,2}^*}{h} \left[1 - \left(\frac{\xi_{1,1,2}^*}{h} \right)^2 \right] \right| < 0,4$$

ist.

b) Bei Annäherung der (unter der Voraussetzung $f'(x) \neq 0$ in $x_1 \leq x \leq x_2$ eindeutigen) Umkehrfunktion $x = g(y)$ von $y = f(x)$ durch Polynome $x = q_{2,1,1}(y)$ bzw. $x = q_{1,1,2}(y)$, welche durch die Punkte $(x_1, f(x_1))$, $(x_m, f(x_m))$, $(x_2, f(x_2))$ gehen sowie $x = g(y)$ in $(x_1, f(x_1))$ bzw. $(x_2, f(x_2))$ in 1.0. berühren, erhält man (vgl. hierzu auch Abschn. 4) die Näherungen

$$x_{II} = x_m - \frac{f_m}{f'_+} \cdot \frac{1}{(f_m - f_+)^2 (f_+ - f_-)^2 (f_- - f_m)} \cdot \begin{vmatrix} 1 & f_m & f_m^2 \\ f_+ + f_- & f_+^2 + f_-^2 & f_+^3 + f_-^3 \\ f_+ - hf'_+ & f_+^2 - 2hf_+f'_+ & f_+^3 - 3hf_+f'_+ \end{vmatrix}$$

⁹⁾ Fr. A. Willers, Z. angew. Math. Mech. 18 (1938), S. 200, Gl. (35).

¹⁰⁾ a. a. O. ¹⁾, Gl. (1) oder a. a. O. ²⁾.

und

$$x_1 = x_m - \frac{f_m}{f'_-} \cdot \frac{1}{(f_m - f_-)^2 (f_- - f_+)^2 (f_+ - f_m)} \cdot \begin{vmatrix} 1 & f_m & f_m^2 \\ f_+ + f_- & f_-^2 + f_+^2 & f_-^3 + f_+^3 \\ f_- + h f'_- & f_-^2 + 2h f_- f'_- & f_-^3 + 3h f_-^2 f'_- \end{vmatrix}$$

x_I, x_{II} schließen x^* ein, wenn in $x_1 \leq x \leq x_2$

$$\frac{d^3 g}{dy^3} = \frac{3f''^2(x) - f'(x)f'''(x)}{f'^5(x)} \text{ monoton}$$

ist.

Für kleine h gilt für die Abweichung $x_{II} - x_I$ der Schranken, $(x^* - x_m = \Theta h, -1 < \Theta < 1)$

$$x_{II} - x_I = \frac{\Theta(1 - \Theta^2)}{12} h^4 \cdot \frac{(15f''^3_* - 10f'_* f''_* f'''_* + f_*'^2 f_*^{(4)})}{f_*'^3} + \dots,$$

wobei $\left| \frac{\Theta(1 - \Theta^2)}{12} \right| < \frac{1}{30}$ ist.

Eingegangen am 2. September 1947.

KLEINE MITTEILUNGEN

Über die Methode von Graeffe.

Bekanntlich besteht die Graeffesche Methode zur Lösung algebraischer Gleichungen darin, aus der gegebenen Gleichung $f(x) = 0$ mit den Wurzeln x_1, \dots, x_n eine andere herzuleiten $f_1(x) = 0$ mit den Wurzeln x_1^2, \dots, x_n^2 , daraus eine neue $f_2(x) = 0$ mit den Wurzeln x_1^4, \dots, x_n^4 usw. und diesen Prozeß der Wurzelquadrirung so lange fortzusetzen, bis die Gleichung $f_k(X) = 0$ bis auf die gewünschte Genauigkeit in die Gleichungen

$$M_1(X) = 0, M_2(X) = 0, \dots, M_j(X) = 0$$

mit den Gradzahlen n_1, n_2, \dots, n_j , wobei $n_1 + \dots + n_j = n$, zerfällt. Die Wurzeln jeder Gleichung M_s haben denselben Modul R_s , so daß die Wurzeln sind

$$X_{s,t} = R_s (\cos A_{s,t} + i \sin A_{s,t}), \quad t = 1, \dots, n_s.$$

Daraus ergibt sich R_s sofort als die n_s -te Wurzel aus dem absoluten Gliede von $M_s = 0$. Ebenso ergibt sich der Modul r_s der zu $X_{s,t}$ gehörigen Wurzel $x_{s,t}$ von $f(x) = 0$:

$$x_{s,t} = r_s (\cos a_{s,t} + i \sin a_{s,t})$$

als

$$r_s = \sqrt[p]{R_s}, \quad \text{wo } p = 2^k.$$

Alle bisherigen Rechenoperationen verlaufen ganz einfach. Unangenehm wird die Rechnung erst für die Bestimmung der Argumente $a_{s,t}$. Und hier teilt sich die Methode in zwei Modifikationen.

I. Die erste Modifikation wurde schon von Graeffe selbst angegeben. Er löste nämlich die Gleichung $M_s(X) = 0$ niemals vollständig, sondern er bestimmte nur den Modul R_s (und daraus r_s). Um jedoch die Argumente $A_{s,t}$ zu finden, löste er die Gleichung

$$f(x' - b) = 0,$$

d. h. er bestimmte wiederum nur die Moduln r'_i ihrer Wurzeln x' auf die oben beschriebene Weise. Um die Argumente zu finden, ging er jetzt folgendermaßen vor.

Er ordnete, die Moduln r_s der Größe nach, dabei jeden Modul r_s so oft hinschreibend, wie seine Vielfachheit n_s beträgt:

$$\varrho_1, \varrho_2, \dots, \varrho_n.$$

Ebenso ordnete er die Moduln r'_i der Größe nach:

$$\varrho'_1, \varrho'_2, \dots, \varrho'_n.$$

Dann bewies er, daß für hinreichend kleines $b > 0$ zwei einander entsprechende Moduln ϱ_i, ϱ'_i zu derselben Wurzel x_i von $f(x) = 0$ gehören, d. h. daß

$$\varrho_i = |x_i|, \quad \varrho'_i = |x_i \mp b|.$$

Dazu muß b so gewählt sein, daß es höchstens gleich der kleinsten Differenz zweier aufeinanderfolgender Moduln von $f(x) = 0$ ist:

$$b \leq \min(r_{i+1} - r_i) \text{ für alle } i.$$

Denn in diesem Falle schneidet jeder Kreis C_s mit dem Mittelpunkt 0 und dem Radius r_s (auf welchem sich n_s Wurzeln von $f = 0$ befinden) nur einen einzigen Kreis C'_k mit dem Radius $x_i + b$ und dem Mittelpunkt $-b$.

Ostrowski¹⁾ bewies dann, daß es nicht einmal notwendig ist, b genügend klein zu wählen. Im Gegenteil ist b willkürlich. Doch muß man dann die Moduln ϱ'_i von $|x_i + b|$ nicht mehr der Größe nach ordnen, sondern so, daß

$$\sum \frac{\varrho_i'^2}{\varrho_i^2} = -2b \frac{c_{n-1}}{c_n} + n + b^2 \sum \varrho_i^{-2}$$

(wo c_n das absolute Glied von $f(x)$ ist), eine Beziehung, die sich einfach beweisen läßt.

Aus ϱ_i' und ϱ_i ergibt sich dann einfach das Argument a_i von x_i .

II. Die zweite Modifikation löst jede Gleichung $M_s(X) = 0$ vollständig, d. h. man bestimmt nicht nur den Modul R_s , sondern auch alle Argumente $A_{s,t}$ der Wurzeln X . Dann aber ist bekanntlich das Argument $a_{s,t}$ von x nicht mehr ohne weiters eindeutig, vielmehr hat man es unter 2^k Werten zu bestimmen. Wie man die Argumente $A_{s,t}$ auf die einfachste Weise bestimmt, werden wir unten angeben. Zur Bestimmung der $a_{s,t}$ aus den $A_{s,t}$ jedoch hat Carvallo bekanntlich folgende Methode angegeben.

Er bestimmt sukzessive (wenn wir der einfacheren Schreibweise halber von jetzt an die Indizes bei den X, x, R, r usw. weglassen)

$$\sqrt[4]{X}, \sqrt[8]{X}, \sqrt[16]{X}, \dots, \sqrt[2^k]{X},$$

indem er jeden gefundenen Wert in die vorhergehende Gleichung der Kette einsetzt. Spaltet man somit

¹⁾ Ostrowski: Recherches sur la méthode de Graeffe et les zéros des polynômes et des séries de Laurent, Acta Mathematica, 72, 1940, p. 242. In der dortigen Formel (88,1) und der vorhergehenden befinden sich zwei Flüchtigkeitsfehler.

das Polynom $f_{k-1}(x)$ in seine Glieder grader und ungerader Ordnung:

$$f_{k-1}(x) = g(x^2) + x h(x^2) = 0,$$

so erfüllt x die Gleichung

$$x = -g(x^2)/h(x^2).$$

Nun ist aber x^2 die Wurzel der vorhergehenden Gleichung $f_k(X) = 0$, die schon vollständig aufgelöst ist. Somit ist

$$x = \sqrt{X} = -g(X)/h(X).$$

D.h. um die Wurzel der vorgehenden Gleichung $f_{k-1} = 0$ zu finden, hat man die Wurzel der Gleichung $f_k = 0$, welche schon bekannt ist, in die Ausdrücke g und h von f_{k-1} einsetzen.

Wenn auch g und h nur etwa die Gradzahlen $n/2$ haben, so ist doch das sukzessive Einsetzen aller Wurzeln von $f_k = 0$ in alle vorgehenden Gleichungen der Kette sehr mühsam. Daran ändert sich auch nichts, wenn man die Rechnungen nur auf 2–3 Dezimalen oder noch weniger ausführt.

Damit ist aber die zweite Modifikation der ersten in ihrer Zweckmäßigkeit weit unterlegen. Und dies war der Grund, daß man den Kunstgriff von Carvallo in seiner bis heute unverändert gebliebenen Form vernachlässigte und die erste Modifikation bevorzugte.

Nun läßt sich aber der Kunstgriff sehr vereinfachen. Zunächst wird man überhaupt das Verfahren nicht schematisch durchführen, sondern sich erst überzeugen, welche Moduln r_s zu reellen Wurzeln gehören. Dies ist schnell festgestellt, wobei man sich verschiedener Hilfsmittel bedienen kann, und die betreffenden Wurzeln scheiden für das Verfahren nach Carvallo aus.

Bei dem Kunstgriff selber hatte man nun bisher nicht beachtet, daß es in Wirklichkeit überhaupt nicht nötig ist, irgendeine Division auszuführen, ja nicht einmal die Funktionswerte $g(X)$ und $h(X)$ zu berechnen. Man hatte nämlich ganz übersehen, daß ja die Wurzel x von $f_{k-1} = 0$ nicht völlig unbekannt ist, sondern daß man nur zwischen zwei bekannten Werten zu wählen hat, nämlich, wenn

$$X = R(\cos A + i \sin A);$$

zwischen

$$x_1 = \sqrt{R}(\cos A/2 + i \sin A/2) \text{ und } x_2 = -x_1.$$

Zu einer solch einfachen Entscheidung braucht man aber nicht $g(X)$ und $h(X)$ zu berechnen und erst recht nicht beide durcheinander zu dividieren. Denn das Argument kann nur zwei Werte annehmen, die sich um π voneinander unterscheiden. Man kann bei der Bestimmung von x also eine Ungenauigkeit begehen, wenn diese nur absolut genommen kleiner ist als π . Nun ist

$$g(X) = \sum c_j R^j \cos jA + i \sum c_j' R^j \sin jA = c e^{i\alpha}$$

$$h(X) = \sum c'' R^m \cos mA + i \sum c''_m R^m \sin mA = c' e^{i\beta},$$

daher

$$x = -g(X)/h(X) = -\frac{c}{c'} e^{i(\alpha - \beta)},$$

d.h. $-x$ hat das Argument $\alpha - \beta$, wo nach obigen Ausführungen α nur bis auf $\pi/2$ genau bestimmt zu werden braucht und ebenso β . Dies ist aber sicherlich der Fall, wenn man die Quadranten kennt, in welchen die Funktionswerte $g(X)$ und $h(X)$ liegen, und dazu wieder brauchen wir von obigen vier Summen nur die Vorzeichen zu kennen. Dies ist aber schnell geschehen, da die Gradzahlen von g und h etwa $n/2$ sind und zudem in beiden Summen von $g(X)$ jedesmal die Faktoren $c_j R^j$ die gleichen sind, und ebenso bei $h(X)$.

Zusammengefaßt haben wir somit folgenden

Satz. Ist $X = R e^{iA}$ die Wurzel der Transformierten $f_1(x)$ von

$$f(x) \equiv g(x^2) + x h(x^2),$$

so bestimmt sich das Argument der Wurzel x von $f(x)$ folgendermaßen:

Man bestimmt die Quadrantennummer q_g von $g(X)$, ebenso q_h von $h(X)$. Dann ist die Quadrantennummer von x entweder

$$q_g - q_h - 1 \quad \text{oder} \quad q_g - q_h - 2.$$

Das Argument von x ist dann derjenige der beiden Werte

$$A/2 \quad \text{oder} \quad \pi + A/2,$$

welcher mit dieser Angabe in Übereinstimmung gebracht werden kann.

Mit dieser Vereinfachung des Kunstgriffes von Carvallo ist die zweite Modifikation der ersten gleichwertig.

Es handelt sich also jetzt noch darum, für die vollständige Lösung jeder Gleichung $M(X) = 0$ von $f_k(X) = 0$ ein zweckmäßiges Verfahren anzugeben. Dies habe ich in einer früheren Arbeit getan²⁾, wovon ich hier nur die Resultate anführe.

Sei

$$M(X) = C_0 X^m + \dots + C_m = 0.$$

Dann verfährt man folgendermaßen:

1. Der Modul R aller Wurzeln von $M = 0$ ergibt sich aus

$$R^m = |C_m/C_0|.$$

Allgemein ist, wie man leicht beweist, der Quotient zweier symmetrischer C absolut, genommen gleich einer Potenz von R :

$$R^{m-2r} = |C_{m-r}/C_r|,$$

so daß, wenn m ungrade ist, R sich einfach als Quotient zweier mittlerer Glieder ergibt:

$$R = |C_{s+1}/C_s|, \quad \text{wo } m = 2s + 1.$$

2. Nachdem man somit R bestimmt hat, setze man $X = RY$, wo $Y = e^{iW}$ unbekannt ist. Setzt man diesen Ausdruck von X in $M(X) = 0$ ein, so erhält man, worauf man bisher noch nie geachtet hatte, eine reziproke Gleichung $M'(Y) = 0$ in Y , weil mit e^{iW} auch $e^{-iW} = Y^{-1}$ eine Wurzel ist.

3. Man dividiert $M'(Y)$ so oft durch $Y^{-1} - 1$ wie dies möglich ist, etwa m' mal. Dann sind m' Wurzeln von $M(X) = 0$ gleich R .

4. Der Quotient $Q(Y) = M'(Y)/(Y^{-1} - 1)^{m'}$ ist ebenfalls ein reziprokes Polynom.

Setzt man daher $Y + Y^{-1} = Z = 2 \cos W$ in $Q(Y)$ ein, so bekommt man eine Gleichung $Q'(Z) = 0$ vom Grade $m'' = (m - m')/2$, deren Wurzeln alle reell und kleiner als 2 sind (absolut genommen). Diese Gleichung läßt sich also einfach nach Graeffe lösen und möge die Wurzeln $Z_1, Z_2, \dots, Z_{m''}$ liefern.

5. $\cos W_i = Z_i/2$ liefert dann m'' Werte für W , also $2m''$ Werte von Y .

6. Diese $2m''$ Werte liefern zusammen mit der m' -fachen Wurzel $X = R$ das vollständige System $X = RY$ der Wurzeln von $M(X) = 0$.

Diese Methode der Lösung von $M(X) = 0$ und im Anschluß daran von $f(x) = 0$ versagt in keinem noch so komplizierten Falle.

Man hat zahlreiche Kunstgriffe vorgeschlagen, um die Konvergenz des Graeffeschen Ver-

²⁾ E. Bode wig: On Graeffe's Method for Solving Algebraic Equations. Quarterly Applied Mathematics, USA., vol. IV, 1946, p. 177–190.

fahrens in komplizierten Fällen, z. B. bei Wurzeln mit benachbarten Moduln oder gar bei nahe beieinanderliegenden Wurzeln, zu beschleunigen. Von Ausnahmefällen abgesehen sind aber alle diese Kunstgriffe unzweckmäßig. Denn bevor man jene Informationen über die Wurzeln hat, hat man — von Ausnahmefällen abgesehen — mehrere Graeffe-Transformationen ausführen müssen und hat erst aus der langsamen Aufspaltung der Gleichungen $f_k(x) = 0$ in die Gleichungen $M_1 = 0, \dots, M_j = 0$ Schlüsse in bezug auf die Wurzeln x ziehen können. Jetzt aber wäre es sehr unzweckmäßig, die geleistete Arbeit beiseite zu schieben und erst $f(x) = 0$ zu transformieren und auf die transformierte Gleichung die Graeffe-Transformationen anzuwenden. Denn man brauchte nur die abgebrochene Gleichungskette noch um einige transformierten Gleichungen zu verlängern, um die neue Gleichung $f_r(X) = 0$ sich völlig aufspalten zu sehen.

Wie ich nämlich in genannter Arbeit bewiesen habe, konvergiert die Folge der Werte

$$\frac{2^k}{\sqrt{R_s}}$$

für $k = 1, 2, \dots$ und festes s quadratisch gegen r_s . Dabei verstehe ich unter quadratischer Konvergenz einer Folge x_1, x_2, \dots mit dem Grenzwert X , daß

$$\lim \frac{X - x_{n+1}}{(X - x_n)^2} = c \neq 0.$$

Zusammenfassung.

1. Dieschematische Ausführung der Graeffe-Transformationen bis zur Aufspaltung der transformierten Gleichung ist allen Kunstgriffen überlegen, da diese Methode quadratisch-konvergente Folgen liefert.

2. Der oben gezeigte Weg der Lösung jeder Gleichung $M_i(X) = 0$ gefolgt von dem vereinfachten Kunstgriff von Carvallo ist die einzig sichere, nie versagende und dabei vorteilhafteste Methode.

Den Haag.

Bode wig.

Beitrag zur inversen Interpolation.

Wir gehen aus von den Identitäten¹⁾

$$\left. \begin{aligned} x = x_0 + (y^* - y_0) [x_0 x_1] \\ - (y - y_0) (y_1 - y) [x_0 x_1 x_2] \\ + (y - y_0) (y_1 - y) (y_2 - y) [x_0 x_1 x_2] \end{aligned} \right\} \quad (1)$$

und

$$\left. \begin{aligned} x = x_0 + (y - y_0) [x_0 x_1] \\ - (y - y_0) (y_1 - y) [x_{-1} x_0 x_1] \\ - (y - y_{-1}) (y - y_0) (y_1 - y) [x_{-1} x_0 x_1] \end{aligned} \right\} \quad (2).$$

Hier ist $y_{-1} < y_0 < y < y_1 < y_2$ angenommen. Wegen

$$[x_0 x_1 x_2] = \frac{1}{3!} \left(\frac{d^3 x}{dy^3} \right)_{\eta_1}, \quad y_0 < \eta_1 < y_2$$

$$[x_{-1} x_0 x_1] = \frac{1}{3!} \left(\frac{d^3 x}{dy^3} \right)_{\eta_2}, \quad y_{-1} < \eta_2 < y_1$$

erkennt man, daß die aus den Gleichungen (1) und (2) folgenden Approximationspolynome 2. Grades den gesuchten Wert einschließen (vorausgesetzt, daß der dritte Differentialquotient im Intervall y_{-1}, y_2 das Vorzeichen nicht wechselt). Ausrechnung und Umformung ergibt dann für die inverse quadratische Interpolation die Einschließungsformeln²⁾

¹⁾ Hier bedeuten

$$[x_0 x_1] = \frac{x_1 - x_0}{y_1 - y_0} \text{ usw.}$$

die „Dividierten Differenzen“. In meiner Schreibweise ist also

$$[y_0 y_1] = \frac{y_1 - y_0}{x_1 - x_0} \text{ usw.}$$

im Gegensatz zu der sonst üblichen.

²⁾ Diese Formeln wurden vom Verfasser bereits auf anderem Wege abgeleitet (vgl. diese Zeitschrift Bd. 28 (1948), S. 122).

$$x = x_0 + \omega \frac{y - y_0}{y_1 - y_0} \left\{ 1 - \frac{y_1 - y}{y_2 - y_1} \left[2 \frac{y_1 - y_0}{y_2 - y_0} - 1 \right] \right\} \quad (3)$$

und

$$x = x_0 + \omega \frac{y - y_0}{y_1 - y_0} \left\{ 1 - \frac{y_1 - y}{y_1 - y_{-1}} \left[1 - \frac{y_1 - y_0}{y_0 - y_{-1}} \right] \right\} \quad (4).$$

Hier bedeutet $\omega = x_1 - x_0$ den konstanten Argument-schritt der Tafel.

Für die Anwendung ist es vorteilhaft, diese Formeln auf eine andere Gestalt zu bringen. Dazu setzen wir

$$\frac{1}{y_2 - y_1} \left[2 \frac{y_1 - y_0}{y_2 - y_0} - 1 \right] \approx \frac{1}{2} \frac{\Delta y_0 - \Delta y_1}{\Delta y_0 \Delta y_1},$$

$$\frac{1}{y_1 - y_{-1}} \left[1 - \frac{y_1 - y_0}{y_0 - y_{-1}} \right] \approx \frac{1}{2} \frac{\Delta y_{-1} - \Delta y_0}{\Delta y_{-1} \Delta y_0}$$

und erhalten so die einfacheren Einschließungsformeln

$$x = x_0 + \omega \frac{y - y_0}{\Delta y_0} \left\{ 1 - \frac{y_1 - y}{2} \frac{\Delta y_0 - \Delta y_1}{\Delta y_0 \Delta y_1} \right\} \quad (5)$$

und

$$x = x_0 + \omega \frac{y - y_0}{\Delta y_0} \left\{ 1 - \frac{y_1 - y}{2} \frac{\Delta y_{-1} - \Delta y_0}{\Delta y_{-1} \Delta y_0} \right\} \quad (6).$$

Als Beispiel soll

$$\frac{\pi}{4} = \arcsin \frac{1}{2} \sqrt{2} \left(\frac{1}{2} \sqrt{2} = 0,70710\,67812 \right) \cdot$$

berechnet werden. Mit den Funktionswerten

$$\begin{aligned} \sin 0,784 &= 0,70611\,74395 & \sin 0,786 &= 0,70753\,22158 \\ \sin 0,785 &= 0,70682\,51811 & \sin 0,787 &= 0,70823\,85430 \end{aligned}$$

ergeben Formel (5) und (6)

$$0,78539\,81633 < \frac{\pi}{4} < 0,78539\,81636.$$

Der Mittelwert ist bis zur 10. Dezimalstelle richtig.

Die Formel für den Mittelwert von (5) und (6) lautet

$$x = x_0 + \omega \frac{y - y_0}{\Delta y_0} \left\{ 1 - \frac{y_1 - y}{4} \frac{\Delta y_{-1} - \Delta y_1}{\Delta y_{-1} \Delta y_1} \right\} \quad (7).$$

Durch Anwendung der Näherungsgleichung

$$\begin{aligned} \frac{(y - y_0) \Delta y_1}{(y - y_0) \Delta y_1 + (y_1 - y) \Delta y_0} &\approx \frac{y - y_0}{\Delta y_0} \\ - \frac{(y - y_0) (y_1 - y) (\Delta y_0 - \Delta y_1)}{\Delta y_0 \Delta y_0 \Delta y_1} \end{aligned}$$

können wir Formel (5) schreiben

$$\begin{aligned} x = x_0 + \omega \frac{(y - y_0) \Delta y_1}{(y - y_0) \Delta y_1 + (y_1 - y) \Delta y_0} \\ + \omega \frac{y - y_0}{\Delta y_0} \frac{y_1 - y}{2} \frac{\Delta y_0 - \Delta y_1}{\Delta y_0 \Delta y_1}, \end{aligned}$$

d. h. die inverse lineare Interpolationsformel

$$x = x_0 + \omega \frac{(y - y_0) \Delta y_1}{(y - y_0) \Delta y_1 + (y_1 - y) \Delta y_0} \quad (8)$$

ergibt (bis auf Glieder 3. Ordnung) einen entgegengesetzten gleichen Fehler wie die Regula falsi

$$x = x_0 + \frac{y - y_0}{\Delta y_0}$$

und eignet sich daher vorzüglich zur Fehlerabgrenzung der inversen linearen Interpolation.

Vertauschen wir in Gl. (1) x_2 und y_2 mit x_0 und y_0 , so folgt

$$\left. \begin{aligned} x &= x_0 + (y - y_0) \frac{[x_0 x_1]}{(y_1 - y_0)(y_1 - y) [x_0 x_0 x_1]} \\ &\quad - (y - y_0)^2 \frac{[x_0 x_1]}{(y_1 - y_0)^2 (y_1 - y) [x_0 x_0 x_1]} \end{aligned} \right\} \quad (9)$$

und wegen

$$[x_0 x_0 x_1] = \frac{1}{y_1 - y_0} \left\{ [x_0 x_1] - \left(\frac{dx}{dy} \right)_0 \right\}$$

ergibt sich das Approximationspolynom 2. Grades

$$x = x_0 + \frac{y - y_0}{y_1 - y_0} \left\{ \frac{y - y_0}{y_1 - y_0} \omega + (y_1 - y) \frac{1}{y'_0} \right\} \quad (10).$$

Vertauschen wir dagegen in Gl. (1) x_2 und y_2 mit x_1 und y_1 , so folgt

$$\left. \begin{aligned} x &= x_0 + (y - y_0) \frac{[x_0 x_1]}{(y_1 - y_0)(y_1 - y) [x_0 x_1 x_1]} \\ &\quad - (y - y_0)^2 \frac{[x_0 x_1]}{(y_1 - y_0)^2 (y_1 - y) [x_0 x_1 x_1]} \end{aligned} \right\} \quad (11)$$

oder, wegen

$$[x_0 x_1 x_1] = \frac{1}{y_1 - y_0} \left\{ \left(\frac{dx}{dy} \right)_1 - [x_0 x_1] \right\}$$

$$x = x_0 + \frac{y - y_0}{y_1 - y_0} \left\{ \left(2 - \frac{y - y_0}{y_1 - y_0} \right) \omega - (y_1 - y) \frac{1}{y'_1} \right\} \quad (12).$$

Aus (9) und (11) erkennt man, daß die Approximationspolynome (10) und (12) den gesuchten Wert einschließen. Für ihren Mittelwert erhalten wir die nicht nur genauere, sondern auch einfachere Formel

$$x = x_0 + \frac{y - y_0}{y_1 - y_0} \left\{ \omega + \frac{1}{2} (y_1 - y) \frac{y'_1 - y'_0}{y'_0 y'_1} \right\} \quad (13).$$

Ihre Genauigkeit wird bei $\omega = 10^{-3}$ im allgemeinen bis zur 10. Dezimalen reichen.

Als Beispiel soll wieder

$$\frac{\pi}{4} = \arcsin \frac{1}{2} \sqrt{2} \left(\frac{1}{2} \sqrt{2} = 0,7071067812 \right)$$

berechnet werden. Mit den Tafelwerten

$$\begin{aligned} \sin 0,785 &= 0,7068251811 & \cos 0,785 &= 0,707382692 \\ \sin 0,786 &= 0,7075322158 & \cos 0,786 &= 0,7066810904 \end{aligned}$$

ergibt unsere Formel bis zur letzten Stelle richtig

$$\frac{\pi}{4} = 0,7853981634.$$

Gl. (13) entsteht auch auf folgende Weise: In der o. a. Arbeit des Verfassers wurde die Interpolationsformel

$$f(x) = (1 - u) y_0 + u y_1 + \frac{1}{2} \omega u (1 - u) [y'_0 - y'_1]$$

abgeleitet. Mit

$$u = \frac{x - x_0}{x_1 - x_0}$$

können wir dieser Gleichung die Form geben.

$$y = y_0 + \frac{x - x_0}{x_1 - x_0} \left\{ (y_1 - y_0) + \frac{1}{2} (x_1 - x) [y'_0 - y'_1] \right\}$$

und hieraus folgt nach Vertauschung von x mit y sofort die neue Formel für die inverse quadratische Interpolation.

Bei dieser Gelegenheit sei auch darauf aufmerksam gemacht, daß die vom Verfasser früher angegebene Formel für die inverse lineare Interpolation der Exponentialfunktion^{*)}

$$x = x_0 + (y e^{-x_0} - 1)$$

das erste Glied ist einer allgemeinen Entwicklung nach Potenzen des Klammerausdrucks. Das ergibt sich sofort aus der Potenzreihe

$$\ln(1 + x) = \sum_{s=1}^{\infty} (-1)^{s+1} \frac{x^s}{s},$$

wenn man x durch

$$\frac{y}{y_0} - 1$$

ersetzt. Für $y = e^x$, $y_0 = e^{x_0}$ entsteht dann die Entwicklung

$$x = x_0 + (y e^{-x_0} - 1) - \frac{1}{2} (y e^{-x_0} - 1)^2 + \dots \quad (14).$$

Berücksichtigt man nur die beiden ersten Glieder, so erhält man für die inverse quadratische Interpolation der Exponentialfunktion die Formel

$$x = x_0 + \frac{1}{2} (y e^{-x_0} - 1) \{ 2 - (y e^{-x_0} - 1) \} \quad (15).$$

Der Fehler ist kleiner als

$$\frac{1}{3} (y e^{-x_0} - 1)^3.$$

Göttingen.

F. K. Rubbert.

^{*)} Formel (2,3) in der o. a. Arbeit des Verfassers, die dort aber einen Druckfehler enthält, indem es nicht

$$e^{-x}, \text{ sondern } e^{-x_0}$$

heißt muß. Entsprechend ist (2,4) zu berichtigen.

BUCHBESPRECHUNGEN

Dr. Wolfgang Haack, Differentialgeometrie. Teil I. (Bücher der Mathematik und Naturwissenschaften, herausgegeben von Dr. Henry Poltz.) 136 S. m. 19 Abb. Wolfenbüttel und Hannover 1948. Wolfenbütteler Verlagsanstalt G. m. b. H. (Notdruck) Preis brosch. 9,— DM.

Nach einer sehr knappen, vom Invariantenbegriff ausgehenden Einführung in die Vektoralgebra behandelt der Verfasser die Raumkurven, die Elemente der Flächentheorie (Fundamentalformen, Sätze von Meusnier und Euler, Gauß' theorema egregium), die Abbildung zweier Flächen aufeinander (längen-, winkel- und flächentreue Abbildung), die Ableitungsformeln von Gauß und Weingarten und die Verträglichkeitsbedingungen von Mainardi und Codazzi, die Geometrie auf der Fläche (geodätische Linien, Parallelismus von Levi-Civita, Gauß-Bonnetsche Integralformel, Poincarés Modell der hyperbolischen Ebene) und schließlich die Minimalflächen. Die Fülle des auf 133 Seiten zusammengedrückten Stoffes bedingt eine erhebliche Kürze der Darstellung, die dem Anfänger vielleicht zu schaf-

fen machen wird, obwohl sich das didaktische Geschick des Verfassers auch hier wieder bestens bewährt. Er knüpft seine Betrachtungen, wie es in den einführenden Vorlesungen über Differentialgeometrie üblich ist, an die anschaulichen geometrischen Sachverhalte an, versteht es aber trotzdem, den Leser schon an den invariantentheoretischen Kern der Probleme heranzuführen. Erfreulich sind die zahlreichen historischen Anmerkungen und die Ausführlichkeit, mit der der Verfasser im Hinblick auf die Anwendungen in der Kartographie die Abbildungen der Flächen aufeinander behandelt. Mißfallen hat dem Rezensenten nur, daß der Verfasser dem Einbruch des Zeitungsjargons in die wissenschaftliche Literatur Vorschub leistet, indem er von „Bertrand-Kurven“ und „Darboux-Vektor“ spricht und eine Funktion „nach Taylor“ entwickelt. Bei seinen weiteren Arbeiten möge der Verfasser beherzigen, was in Wustmanns Sprachdummheiten über dieses Thema gesagt wird! Dann sei noch bemerkt, daß man bei der Taylorsche Entwicklung eines Vektors nicht von einem Zwischenwert θ in der Restgliedformel spre-

chen kann, weil die θ -Werte in den Restgliedformeln der einzelnen Komponenten in der Regel nicht übereinstimmen werden, und schließlich muß festgestellt werden, daß die Huygenssche Schleppkurve nicht dasselbe ist wie die Verfolgungskurve — „Hundekurve“ nennt sie der Verfasser! Das Buch zeigt bei aller zeitbedingten Schlichtheit der Aufmachung doch ein vorbildlich klares Satzbild und vorzügliche Figuren. Alles in allem eine sehr erfreuliche Neuerscheinung.

Freiberg.

Grüss.

Dr.-Ing. habil. H. Athen, Vektorrechnung (Bücher der Mathematik und Naturwissenschaften, herausgegeben von Dr. Henry Poltz). 90 S. m. 60 Abb. Wolfenbüttel-Hannover 1948. Wolfenbütteler Verlagsanstalt GmbH. (Notdruck.) Preis brosch. 6,50 DM.

Das Bändchen will eine Einführung in die Vektorrechnung geben. In der Darstellung schließt es sich an die bekannten größeren Werke an, benutzt dabei aber die ältere Schreibweise. (Für skalares Produkt runde, für Vektorprodukt eckige Klammern.) Nach Behandlung der Vektoralgebra, der Vektoranalysis und ihrer Anwendung in der Potentialtheorie folgt eine kurze anschauliche Einführung in die Tensorrechnung. Nur wenige Beispiele werden gebracht; denn Verfasser beabsichtigt in einem Ergänzungsband die wichtigsten Anwendungen der Vektor-

rechnung in Mathematik, Physik und Technik zu behandeln.

Das Bändchen gibt eine gute Einführung in das behandelte Gebiet. Besonders gut wird es neben Vorlesungen zu gebrauchen sein. Für das Selbststudium scheint mir die Darstellung an manchen Stellen etwas zu konzentriert zu sein.

Dresden.

Willers.

Karl Daeves und August Beckel, Großzahlforschung und Häufigkeits-Analyse. Ein Leitfaden. 65 S. m. 17 Abb. Weinheim/Bergstr. und Berlin 1948. Verlag Chemie GmbH. Preis brosch. 3,20 DM.

Das Bändchen gibt eine Anleitung zur Durchführung von Großzahluntersuchungen insbesondere unter Verwendung von Häufigkeitspapier und Wahrscheinlichkeitsnetzen, wobei meist logarithmische Merkmalskalen benutzt werden. Dabei handelt es sich darum, Kollektive beliebiger Verteilung in eine kleine Zahl von Teilkollektiven mit Gaußscher Verteilung aufzuspalten. Diese werden durch Mittelwert und durch die Spanne der Klassengrenzen charakterisiert, außerhalb deren beiderseits je 5% der Kollektivgegenstände liegen, also nicht wie üblich durch Mittelwert und Streuung.

Dresden.

Willers.

Die besprochenen und angezeigten Bücher sind durch den Buchhandel zu beziehen.

EINGEGANGENE BÜCHER

Bei der Schriftleitung sind folgende Bücher eingegangen (ausführliche Besprechung bleibt vorbehalten).

Dr.-Ing. Fr. Bošnjaković (o. Prof. an der Universität Zagreb), Technische Thermo-Dynamik (Wärmelehre und Wärmewirtschaft in Einzeldarstellung. Herausgegeben von Dr.-Ing. W. Pauer, Bd. 11, I. Teil.) 3. unveränderte Auflage, XVI + 327 S. mit 249 Abb. und 3 Tafeln. Dresden und Leipzig 1948. Verlag Theodor Steinkopff. Preis: geb. 18,—; brosch. 17,— DM.

Dr.-Ing. Hans Umstätter (Kaiser Wilhelm-Institut für Physikalische Chemie und Elektrochemie), Strukturmechanik. Ein Beitrag zur Phy-

sik der Kolloide. XVI + 237 S. mit 1 Titelbild, 143 Abb. und 34 Tab. Dresden und Leipzig 1948. Verlag Theodor Steinkopff. Preis: gebunden 19,—; brosch. 17,— DM.

Dr.-Ing. Johannes Johannson, Das Cross-Verfahren, Die Berechnung biege-fester Tragwerke nach der Methode des Momentenausgleichs. 123 S. mit 18 Zahlenbeispielen und 137 Abbildungen. Berlin—Göttingen—Heidelberg 1948. Springer-Verlag. Preis: brosch. 14,40 DM.

NACHRICHTEN

Richard Grammel sechzig Jahre alt.

Es ist ein großer Kreis von Menschen, die des Jubilars freudig bewegt gedenken: Die Fachgenossen in aller Welt erkennen den hervorragenden Gelehrten und Forscher schon seit langem als einen der ersten in ihrem Kreise an, eine große Anzahl von Schülern bewundert und verehrt ihren trefflichen Lehrer, die Kollegen und akademischen Mitbürger danken dem tatkräftigen Mann, der ihre Hochschule aus Trümmern neu begründete, seine Freunde lieben den warmherzigen Menschen und schätzen den geistreichen Erzähler und Plauderer.

Richard Grammel hat eine bedeutende Wirksamkeit entfaltet; sie reicht weit über den Kreis seiner eigentlichen Schüler hinaus. Es gibt wohl kaum einen Ingenieur, und gewiß keinen wissenschaftlich interessierten, dem der Name Grammel nicht ein fester Begriff wäre. Dieser Name ist insbesondere verbunden mit drei großen Publikationswerken: Mit der Zeitschrift „Ingenieur-Archiv“, die Grammel gründete, und die er seit nahezu zwanzig Jahren leitet, mit dem „Biezeno-Grammel“, der zur „Bibel“ des Faches geworden ist, und — etwas weniger in die Breite wirkend — mit der Redaktion der drei Mechanikbände des „blauen“ Handbuches der Physik.

Im weiteren wissenschaftlichen Werk des Jubilars beeindruckt den Betrachter die Konsequenz, mit der Grammel die Probleme eines Gebietes aufgreift, zu

Ende denkt und abschließt, wie er dann zu einem Nachbargebiet übergeht und dort dieselbe gründliche und vollständige Arbeit leistet. Die etwa sechzig Schriften, die das Verzeichnis nennt, lassen sich nahezu vollständig in einige wenige Gruppen einordnen: Nach einigen Veröffentlichungen aus verschiedenen Themenkreisen wandte Grammel sich dem Kreisel zu. Es entstanden eine Reihe von wesentlichen Untersuchungen, die ihre Krönung fanden in der Monographie „Der Kreisel“ vom Jahre 1920. Die Untersuchung über „Die kritischen Zustände rasch umlaufender Wellen“ bildete den Übergang vom Kreisel zu den Stabilitätsproblemen, die die Jahre bis etwa 1929 ausfüllen. Dann leitete die Arbeit „Über den Einfluß der Wellentorsion auf die kritische Drehzahl“ über zu der großen Gruppe von Arbeiten aus den dreißiger Jahren über die Torsionsschwingungen der Wellen von Maschinenanlagen, durch die die praktisch so brauchbaren Rechenmethoden geschaffen und in eine handliche Form gebracht wurden, und in denen überdies eine Reihe grundsätzlich neuer Erkenntnisse zutage gefördert wurden, wie sie sich insbesondere um die Stichworte „Torsion erster Art“ und „Torsion zweiter Art“ ranken. Es schließt sich dann die Gruppe der Arbeiten über die Scheiben an.

Zusammengefaßt werden die Problemkreise, denen diese Untersuchungen angehören, in dem schon genannten, gemeinsam mit C. B. Biezeno verfaßten

Werk „Technische Dynamik“, einem Buch, auf dessen mehr als tausend Seiten keine Einzelheit aus fremder Hand einfach übernommen wurde, wo vielmehr jede Tatsache neu aufgefaßt und originell dargestellt ist.

Diese Besonderheit Grammel'scher Arbeitsweise, eine Frage neu zu stellen oder eine alte auf überraschend neue Art zu beantworten, spiegelt sich ferner deutlich wieder in mehreren Arbeiten, von denen als kennzeichnend vor allem aus neuerer Zeit besonders genannt seien: „Kinetisch unbestimmte Systeme“, „Die Erklärung der hohen Sprengfestigkeit umlaufender Scheiben“, „Eine Verallgemeinerung der Kreis- und Hyperbelfunktionen“ und (die noch im Druck befindliche) „Scherprobleme“.

Es ist noch lange nicht Zeit, Rückschau zu halten und eine Würdigung im ganzen zu geben. Die auf ihrer vollen Höhe stehende Schaffenskraft des Jubilars wird uns, so dürfen wir, ohne enttäuscht zu werden, hoffen, noch viele weitere Ergebnisse seines Forschens bringen und uns noch wesentliche Erkenntnisse und Einsichten verschaffen.

Seinen Lebensweg begann Richard Grammel im Schwabenland. Er wurde am 3. März 1889 in Klosterreichenbach im Oberamt (Kreis) Freudenstadt geboren. Seine Studien in Stuttgart, München und Tübingen schloß er im Jahre 1913 mit der Promotion in Tübingen ab. Dann habilitierte er sich 1915 in Danzig, wo er einige Jahre Assistent bei Hans Lorenz war. Aus einer kurzen Tätigkeit an der Universität Halle berief ihn die Technische Hochschule seines Heimatlandes im April 1920 zum Ordinarius für Technische Mechanik und Technische Wärmelehre. In Stuttgart lebt und lehrt Grammel seitdem.

Die Erfolge des Forschers, Lehrers und tatkräftigen Mannes haben eine Reihe von Anerkennungen und Ehrungen gebracht. Den erst Vierzigjährigen erkor seine Hochschule zum Rektor im Jahre ihrer Zentenarfeier. Die Eidgenössische Technische Hochschule in Zürich reichte ihn zugleich unter ihre Ehrendoktoren ein. Nach dem Zusammenbruch des Jahres 1945 berief ihn seine Hochschule erneut, und zwar drei Jahre lang, zum Rektor. Hier führte er den Neuaufbau in so hervorragender Weise durch, daß die Technische Hochschule Stuttgart einen beachtlichen Vorsprung vor ihren Schwesteranstalten erlangte. Man gab dem Dank für diese Leistung Ausdruck durch die Ernennung zum Ehrenbürger der Hochschule. In der Gesellschaft für

angewandte Mathematik und Mechanik (der amerikanisch besetzten Zone) führt Richard Grammel den Vorsitz.

Karlsruhe.

Karl Klotter.

Sowjetischer Erfindungsschutz für deutsche Erfinder

Das Interesse deutscher Erfinder am Erfindungsschutz in der UdSSR. ist durch die Errichtung einer Verwaltung für Patentschutz in der UdSSR. bei der SMAD. in erheblichem Maße angeregt worden. Diese Berliner Verwaltungsstelle, Berlin-Friedrichsfelde, Schloßstr. 32c—33 ist von der zentralen Verwaltung für Patentschutz in Moskau zur Annahme von Schutzanmeldungen und zum Ankauf von Erfindungen ermächtigt worden. Die zahlreichen aus allen Zonen Deutschlands eingehenden Anfragen beweisen, daß mit der Errichtung dieser Verwaltungsstelle und der dadurch gebotenen Überbrückung der sonst unüberwindbaren Valuta-Schwierigkeiten einem dringenden Bedürfnis begegnet ist. Die Erfinder sollten jedoch beachten, daß die sowjetische Industrie einen hohen Entwicklungsstand aufzuweisen hat. Nur Erfindungen, die das berücksichtigen und Neues und Fortschrittliches bieten, haben Aussicht auf praktische Ausnutzung. Jede auf Grund eines Urheberrechtes geschützte Erfindung, die als fortschrittlich und industriell ausnutzbar beurteilt ist, wird von einer besonderen Amtsstelle aus auf tatsächliche und weitgehendste Ausnutzung überwacht. Diese Überwachung dient der schnellen Entwicklung der Industrie, kommt aber gleichzeitig den Erfindern zugute. Da die sowjetische Industrie im letzten Jahrzehnt eine gewaltige Ausdehnung erfahren hat, und da jede Ausnutzung einer in diese Industrie hinein passenden Erfindung einen gesetzlich festgelegten Anspruch auf Entschädigung auslöst, so kann dieser eine erhebliche Höhe erreichen. Eine gesetzliche vorgeschriebene Skala setzt die Entschädigung in Übereinstimmung mit dem durch die Erfindung erzielten Nutzen fest.

Göttingen: Zu ordentlichen Mitgliedern der Akademie der Wissenschaften in Göttingen wurden ernannt: Dr. W. Magnus, o. Prof. der Mathematik an der Universität Göttingen, und Dr. F. Rellich, ebenfalls o. Prof. d. Mathematik an der Universität Göttingen.

Dr. Stellmacher hat sich im Herbst 1948 an der Universität Göttingen für das Fach Mathematik habilitiert.

ZUSCHRIFT AN DEN HERAUSGEBER

Zu G. Plato: Über den Zusammenhang von Schwingungsdauer und Rückstellkraft. Z. angew. Math. Mech. 28 (1948), S. 94.

Ein interessantes Problem ist in dieser Kl. M. angeschnitten. „Gefühlsmäßig“ bin ich mit dem allgemeinen Ergebnis nicht einverstanden. Nimmt man bei einem schwingenden Einmassensysteme an, daß im ersten Intervalle des Maximalausschlages $0 < a < a_1$ die Schwingungsdauer $T_1(a)$ veränderlich ist, so kann für das zweite Intervall $a_1 < a < a_2$ sehr wohl eine konstante Schwingungszeit $T_2(a)$ bestehen. Für das erste Intervall ist das Potential dann keine Funktion von a von der Gestalt ca^2 und für das zweite Intervall auch nicht. Für dieses Intervall ist das Potential durch den Verlauf im ersten Intervall und durch $T_2(a)$ bestimmt.

Das anschaulich gewonnene Ergebnis werde ich in einer getrennten Kleinen Mitteilung beweisen.

Ich überlasse es Herrn Plato, den Fehler in seiner mathematischen Beweiskette zu finden; zum Schlusse kann doch nur einer von uns recht haben.

Schlewecke über Bockenem.

G. Weber.

Erwiderung.

Die mitgeteilte Überlegung beschränkt sich auf den Fall eines durch eine Potenzreihe darstellbaren Potentials, was der Kürze halber zwar nicht ausdrücklich erwähnt aber doch aus dem Text ersichtlich war, da man ja andernfalls nicht von analytischer Fortsetzung sprechen kann. Die Übertragung der Überlegung auf nicht-analytische Potentiale scheitert daran, daß zu einer Folge $V_n(x)$ eine Folge $T_n(a)$ ($n = 1, 2, \dots$) gehören kann, deren Schwankung in einem gegebenen a -Intervall gegen Null strebt.

Darmstadt.

G. Plato.